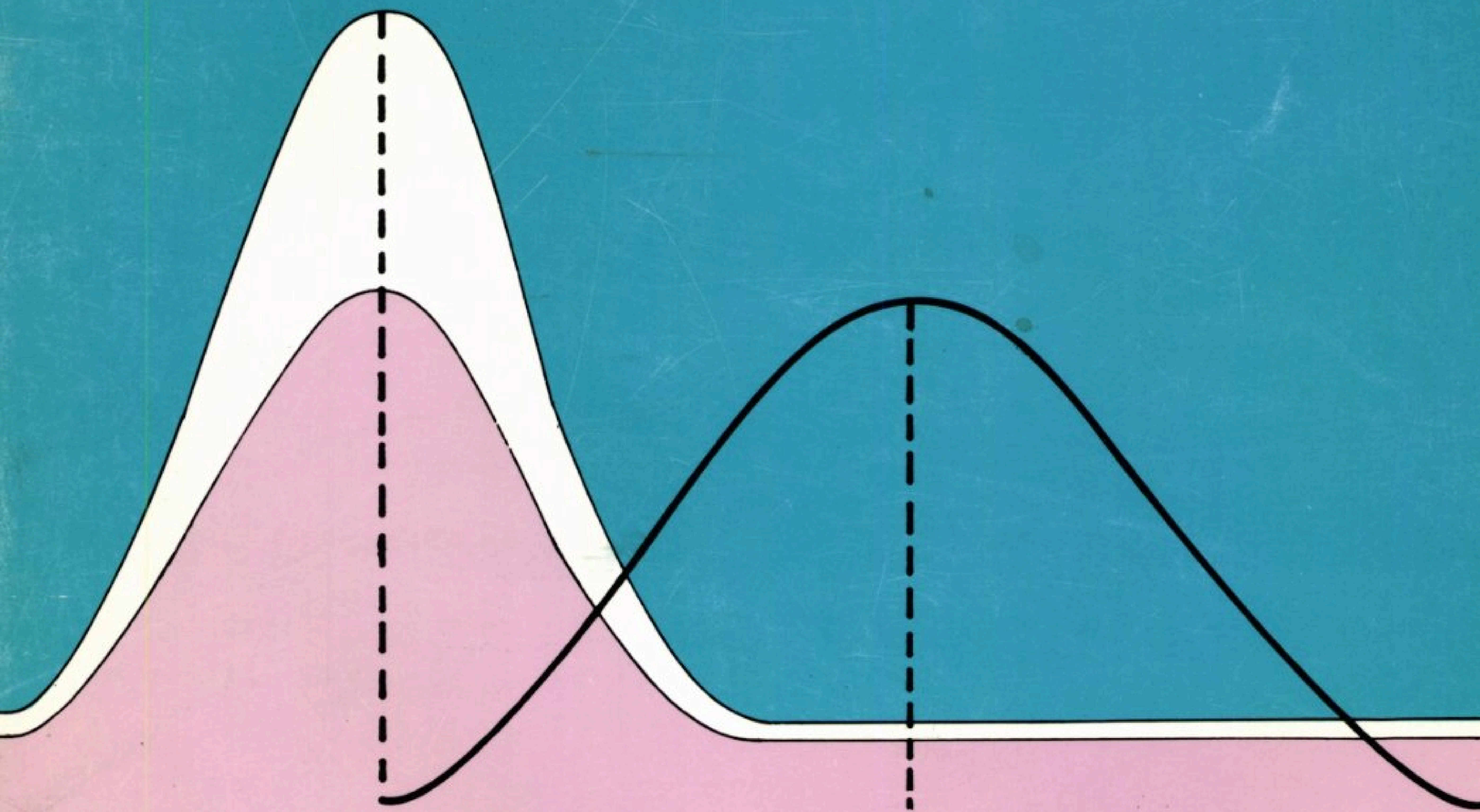
مبادىء
الإحصاء والاحتمالات



الدكتور عدنان بن ماجد عبدالرحمن بري

الدكتور محمود محمد إبراهيم هندي

الدكتور أنور أحمد محمد عبدالله

النشر العلمي والمطابع

جامعة الملك سعود

مبادئ
الإحصاء والاحتمالات

تأليف

الدكتور عدنان بن ماجد عبدالرحمن بري
أستاذ مشارك

الدكتور محمود محمد إبراهيم هندي
أستاذ مشارك

قسم الإحصاء - كلية العلوم - جامعة الملك سعود

الدكتور أنور أحمد محمد عبدالله
قسم الإحصاء - كلية العلوم - جامعة الملك سعود (سابقًا)

إصدار: النشر و المطابع - جامعة الملك سعود
ص.ب. : ٢٤٥٤ - الرياض ١١٤٥١ - المملكة العربية السعودية

© ١٤١٨هـ/١٩٩٧م جامعة الملك سعود

جميع حقوق الطبع محفوظة. غير مسموح بطبع أي جزء من أجزاء هذا الكتاب، أو خزنه في أي نظام لخزن المعلومات واسترجاعها، أو نقله على أية هيئة أو بأية وسيلة سواء كانت إلكترونية أو شرائط ممغنطة أو ميكانيكية، أو استنساخًا، أو تسجيلًا، أو غيرها إلا بإذن كتابي من صاحب حق الطبع.

الطبعة الأولى: ١٤١٢هـ (١٩٩١م)

الطبعة الثانية: ١٤١٥هـ (١٩٩٤م)

الطبعة الثالثة: ١٤١٨هـ (١٩٩٧م)

فهرسة مكتبة الملك فهد الوطنية

بري، عدنان بن ماجد بن عبدالرحمن

مبادىء الإحصاء والاحتمالات/عدنان ماجد بري، محمود محمد إبراهيم هندي، أنور محمد أحمد عبدالله ـ ط٣. ـ الرياض.

٣٩٥ ص، ١٧×٢٤ سم

ردمك ٨ ـ ٥٦٩ ـ ٠٥ ـ ٩٩٦٠

١ ـ الإحصاء ٢ ـ الاحتمالات أ ـ هندي، محمود محمد إبراهيم (م. مشارك) ب ـ عبدالله، أنور أحمد محمد (م. مشارك) جـ ـ العنوان

ديوي ٥١٩ ١٨/٠٣٥٧

رقم الإيداع: ١٨/٠٣٥٧

حكمت هذا الكتاب لجنة متخصصة شكَّلها المجلس العلمي بالجامعة، وقد وافق المجلس على طباعته في اجتماعه الرابع عشر للعام الدراسي ١٤٠٧/١٤٠٨هـ الذي انعقد بتاريخ ١٤٠٨/٥/٧هـ الموافق ١٩٨٧/١٢/٢٧م. ثم وافق المجلس على إعادة طباعته طبعة ثالثة في اجتماعه الثاني عشر للعام الدراسي ١٤١٧/١٤١٨هـ الذي عُقد بتاريخ ١٤١٨/١/١٢هـ الموافق ١٩٩٧/٥/١٨م.

مطابع جامعة الملك سعود ١٤١٨هـ

مقدمة الطبعة الثالثة

نظرًا للإقبال المتزايد على تعلم مبادىء الإحصاء والاحتمالات من قبل الدارسين في شتى مجالات العلوم والمعرفة، جاء هذا الكتاب ليفي بحاجة هؤلاء جميعًا، وبخاصة طلابنا الأعزاء بجامعة الملك سعود وغيرها من الجامعات بالمملكة والوطن العربي الكبير، وليعاونهم في فهم المصطلحات الأساسية لمبادىء الإحصاء والاحتمالات، فقد تضمنت هذه الطبعة أمثلة توضيحية وتمارين. ويهتم هذا الكتاب من الفصل الأول وحتى الفصل الخامس بالتعريف بعلم الإحصاء وتنظيم البيانات جدوليًا وبيانيًا، وحساب مؤشرات خاصة بالنزعة المركزية والتشتت والارتباط والانحدار مع الأمثلة والتمارين لعدد كبير من التطبيقات في العلوم المختلفة. ثم جاءت الفصول الستة التالية من السادس وحتى الحادي عشر لتوضح ماهية علم الاحتمال ومبادئه مثل مسلمات الاحتمال والاستقلال والاحتمال الشرطي ونظرية بييز، مع بعض الأمثلة والتطبيقات والتمارين المتنوعة، ثم الانتقال إلى المتغير العشوائي المنفصل والمتصل، وبعض التوزيعات الخاصة مثل ذي الحدين وفوق الهندسي وبواسون والطبيعي، ثم توزيع المعاينة ثم الاستدلال بشقيه التقدير واختبار الفروض للمتوسطات والنسبة.

وجاءت هذه الطبعة الثالثة بعد إجراء بعض التعديلات الطفيفة لبعض المصطلحات الإحصائية والاحتمالية لزيادة توضيحها وتضمينها ـ كما أسلفنا ـ أمثلة وتمارين إضافية، وكذلك إعادة إخراج بعض الأشكال البيانية باستخدام الحاسوب لزيادة دقتها، وكذلك تصحيح بعض الأخطاء المطبعية التي اكتشفها المؤلفون أو الزملاء المدرسون والمدرسون المساعدون بقسم الإحصاء وبحوث العمليات بكلية العلوم

جامعة الملك سعود، مما يجعل الطبعة الثالثة أكثر اكتمالاً ونفعًا - إن شاء الله - للطلبة والباحثين في مختلف العلوم، والذين ينشدون الإلمام بمبادىء الإحصاء والاحتمالات الأساسية.

والله من وراء القصد وهو يهدي السبيل.

المؤلفـــــون

مقدمة الطبعة الثانية

نظرًا للإقبال المتزايد على تعلم المبادىء الأساسية في علم الإحصاء والاحتمالات من قِبل الدارسين في شتى أنواع العلوم وبخاصة في العلوم الطبية والهندسية والزراعية والإدارية والمالية والاجتماعية والإعلامية . . . إلخ . جاء هذا الكتاب ليفي بحاجة هؤلاء - وبخاصة طلابنا الأعزاء - ومساعدتهم في فهم المصطلحات الأساسية مثل التعريف بالمجتمع الإحصائي والعينة الإحصائية والبيانات الإحصائية والاستمارة الإحصائية مع ذكر الأمثلة التوضيحية والتمارين كما تم شرحه تفصيلاً في الفصل الأول من هذا الكتاب . ثم أتت بعد ذلك الفصول الأربعة - من الثاني إلى الخامس - لتوضح كيفية تنظيم البيانات الإحصائية في جداول إحصائية ورسوم بيانية والاستفادة من هذا التنظيم الجدولي والبياني في حساب مقاييس النزعة المركزية ومقاييس التشتت والتفلطح والالتواء وحساب معاملي الارتباط والانحدار مع التوضيح بالشرح والأمثلة والتمارين المتنوعة لعدد كبير من التطبيقات في العلوم المختلفة. ثم جاءت الفصول الخمسة التالية - من السادس إلى العاشر - لتوضح ماهية علم الاحتمال ومبادئه ومصطلحاته الأساسية، مثل : التجربة العشوائية وفضاء العينة والحدث وتعريف الاحتمال التقليدي والاحتمال النسبي والتعريف الحديث للاحتمال (مسلمات الاحتمال) والنتائج الأساسية لمسلمات الاحتمال وتعريف الاحتمال الشرطي والاستقلال للأحداث وبعض التطبيقات والأمثلة والتمارين المختلفة، ثم الانتقال إلى تعريف المتغير العشوائي المنفصل والمتصل وبعض التوزيعات الإحصائية للمتغير المنفصل، مثل : توزيع ذي الحدين وتوزيع فوق الهندسي وتوزيع بواسون، وكذلك التوزيعات الإحصائية للمتغير المتصل، مثل :

التـوزيع الـطبيعي وتـوزيع ت وتقريب توزيع ذي الحدين للطبيعي وتوزيع المعاينة للمتـوسـطات والنسبـة ونـظرية النهـاية المركزية، مع ذكر بعض التطبيقات والأمثلة والتمارين المتنوعة.

ثم اختتم الكتاب بالفصل الحادي عشر بعرض مبسط للاستدلال الإحصائي بشـطريه، التقدير (بنقطة وبفترة الثقة)، واختبارات الفروض (للمتوسطات والنسبة). وذلك بالتوضيح بالأمثلة المبسطة والتمارين المتنوعة في مجالات العلوم المختلفة، والتي تساعد الدارس من أبنائنا الطلاب بجامعة الملك سعود وغيرهم بالجامعات الأخرى بالمملكـة العـربية السعودية وفي الوطن العربي، مما لمسه المؤلفون أثناء توزيع الطبعة الأولى قبل نفادها.

وجـاءت هذه الـطبعـة الثـانية بعـد إجراء بعض التعديلات الطفيفة لبعض المصطلحات الإحصائية لزيادة توضيحها، وكذلك تصحيح بعض الأخطاء المطبعية التي اكتشفها المؤلفون أو الدارسون من الطلاب أو التي اكتشفها المدرسون المساعدون بقسم الإحصاء بكلية العلوم جامعة الملك سعود. مما جعل هذه الطبعة الثانية أكثر اكتمالًا ونفعًا إن شاء الله للطلبة الدارسين والباحثين في مختلف العلوم، الذين ينشدون الإلمام بمبادىء الإحصاء والاحتمالات الأساسية.

والله من وراء القصد وهو يهدي السبيل.

المؤلفون

الطبعة الثانية ١٤١٥هـ (١٩٩٤م).

مقدمة الطبعة الأولى

أصبح لعلم الإحصاء في الوقت الحاضر دور كبير في جميع العلوم مثل الفيزياء والكيمياء والأحياء والإلكترونيات والاتصالات والزراعة والاقتصاد والتعليم وعلم النفس والعلوم الطبية وغيرها من العلوم والهندسة. وربما اعتمدت هذه العلوم على تحليلات مادة الإحصاء في تقدمها المطّرد. والهدف من هذا الكتاب تقديم بعض الأسس العامة لمبادىء علم الإحصاء والاحتمالات. ويبدأ كل فصل بمقدمة عن الهدف ثم بعد هذه التعاريف والنظريات والأمثلة التي توضح الموضوعات المستهدفة بكل فصل. أما الفصول الأولى من الكتاب فتتناول طرق عرض التوزيعات التكرارية وتلخيصها بالطرق العددية والبيانية وما يتعلق بها من مقاييس: النزعة المركزية، والتشتت والالتواء، والتفرطح، وكذلك أشكال الانتشار لمتغيرين. وما يؤدي من دراسة الارتباط والانحدار. وفي الفصول الأخيرة يتطرق الكتاب لمناقشة بعض مبادىء نظرية المجموعات وطرق العدّ ومبادىء الاحتمالات والمتغيرات العشوائية وبعض التوزيعات الإحصائية وتطبيقاتها وتوزيع المعاينة واستخدامه في فترات الثقة واختبار الفرضيات والمعنويات.

وجاء هذا الكتاب باللغة العربية في كل محتوياته ما عدا المعادلات فقد وضعناها بالرموز اللاتينية. ولذلك فهي تقرأ باستمرار من اليسار إلى اليمين ولا يَعُدّ ذلك قصورًا في اللغة العربية إذ بالإمكان كتابة هذه المعادلات بالرموز العربية دون أية صعوبة. ولكن الهدف من هذا الإجراء التيسير على القارىء في الاستفادة من المراجع الأجنبية.

كما استعنا في إعداد الكتاب بكثير من المراجع العربية والأجنبية المسرودة أسماؤها في نهاية الكتاب وكذلك بخبرتنا في التدريس بجامعة الملك سعود. وقد رأينا في وضعنا لمادة هذا الكتاب أن يكون صالحًا وكافيًا حسب تطور المناهج في جامعة الملك سعود لطلاب المستوى الأول من كلية العلوم أو من هم في مستواهم.

وأخيراً نتوجه بالشكر لكل من ساعد على تهيئة المناخ المناسب لإبراز هذا الكتاب بصورته الحالية والله ولي التوفيق.

المؤلفون

المحتويات

الصفحة

الفصـل الرابع : مقاييس التشتت

الفصل الخامس: الارتباط والانحدار

الفصل السادس : المجـمـوعـات

الفصل السابع : طرق العَدّ

الفصل الثامن : مباديء الاحتمالات

الجداول

١

مقدمة في علم الإحصاء والاحتمالات

● نبذة عن علم الإحصاء ● تعريف علم الإحصاء ● المجتمع الإحصائي والعينة الإحصائية ● البيانات ● المعلمة والإحصائية ● المتغير ● مصادر جمع البيانات الإحصائية ● الاستمارة الإحصائية ● تمارين

(١ - ١) نبذة عن علم الإحصاء

نشأ علم الإحصاء في العصور الوسطى لاهتمام الدول بتعداد أفراد المجتمع حتى تتمكن كل دولة من تكوين جيش قوي يستطيع الدفاع عن حدودها إذا وقع عليها اعتداء من إحدى الدول أو هاجمت دولة ما طمعًا في التوسع والثروة. وكذلك اهتمت الدول بحصر ثروات الأفراد حتى تتمكن من فرض الضرائب وتجميع الأموال اللازمة لتمويل الجيش وإدارة شئون البلاد. ثم توسعت عمليات التعداد والحصر لتشمل بيانات عن المواليد والوفيات والإنتاج والاستهلاك. وبذلك نشأت الحاجة إلى تنظيم هذه البيانات وتلخيصها ووضعها في صورة جداول أو رسم بياني أو تصويري حتى يسهل الرجوع إليها والاستفادة منها بأسرع وقت ممكن وقد أطلق على هذه الطرق «علم الدولة أو علم الملوك ثم علم الإحصاء».

وكلمة "statistics" مشتقة من كلمة "status" وتعني الدولة باللاتينية أو كلمة "statista" بالإيطالية وتعني الدولة أيضًا. هذا كل ما كان يعرف عن علم الإحصاء في ذلك الوقت. حيث كان التحليل الإحصائي للوصول إلى نتائج تستخدم في اتخاذ

القـرارات من الأشياء التي لم تستخـدم بعـد، رغم أنـه قد ظهـرت الحـاجة العامة لاستخدامها ولا سيما بعد تطور علم الاحتمالات في القرنين السابع عشر والثامن عشر الميلاديين بفضـل جهـود العلماء: باسكال (Pascal) وبرنولي (Bernoulli) ودي موافر (de Moivre) ولابلاس (Laplace) وجاوس (Gauss).

ظل الاعتقاد في ذلك الوقت أن علم الإحصاء هو العلم الذي يختص بالطرق العلمية لجمع وتنظيم وعرض البيانات إما في صورة بيانية أو جدولية. حتى إن بعض الأشخاص قليلي الاطلاع ومحدودي التعليم في الوقت الحاضر يعتقدون أن الإحصاء ما هو إلا هذه الطرق. إلا أنه بعد التطور أصبحت الحاجة ملحة إلى تحليل البيانات التي جمعت كالتنبؤ بعدد السكان بعد فترة زمنية بناءً على التعدادات الموجودة، أو التنبؤ بالإنتاج والاستهلاك، أو طرق أخذ العينات وتصميم التجارب. وقد ساعد على ذلك تطور علم الاحتمالات الذي كان له دور كبير في تحليل البيانات واتخاذ القرارات المناسبة بناء على هذا التحليل.

وقـد امتد التطبيق الإحصائي إلى مجالات العلوم الأخرى: كالطب والزراعة والفيزياء، وفي القرن العشرين كثرت الحاسبات الإلكترونية وتنوعت أحجامها وقدرتها ودقتها. الأمر الذي ساعد على تقدم علم الإحصاء بشكل كبير.

وفي الآونة الأخيرة لاحظنا أن معظم الأبحاث الأكاديمية في علم الإحصاء استخدم أصحابها الحـاسبات؛ إما في إتمام البحث ذاته أو التطبيق العددي للنتائج التي حصلوا عليها. كما أن هناك مشروع تعريب البرامج الإحصائية بقسم الإحصاء بالجامعة.

(١ - ٢) تعريف علم الإحصاء

(Statistics)

يعرف علم الإحصاء بأنه: ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وتلخيصها وعرضها وتحليلها وذلك للوصول إلى نتائج مقبولة وقرارات سليمة على ضوء هذا التحليل.

وسوف نتناول بعون الله تعالى كل طريقة بالشرح المفصل والأمثلة التوضيحية.

ينقسم علم الإحصاء إلى قسمين أساسيين هما الإحصاء الوصفي (descriptive or deductive statistics) والإحصاء الاستدلالي (inductive statistics or statistical inference).

فالإحصاء الوصفي: هو طرق تنظيم المعلومات وتلخيصها. والغرض من التنظيم هو المساعدة على فهم المعلومات. والطرق الوصفية تحتوي على توزيعات تكرارية (الجداول التكرارية)، ورسوم بيانية، وطرق حساب مقاييس النزعة المركزية (المتوسطات)، ومقاييس التشتت ومختلف القياسات الأخرى.

والإحصاء الاستدلالي: هو الوسائل العلمية التي تجرى لسبر معالم المجتمع بناءً على المعلومات التي تم الحصول عليها من العيّنة المأخوذة منه وفق الطرق الإحصائية المعلومة.

(١ ـ ٣) المجتمع الإحصائي (Population) والعيّنة الإحصائية (Sample)

يعرف المجتمع بأنه مجموعة ذات خصائص مشتركة من الأفراد محل الدراسة.
ويقسم المجتمع الإحصائي إلى قسمين:

١ ـ **محدود**: وهو الذي يكون فيه عدد محدود من الأفراد مثل عدد حبات الطماطم في صندوق، عدد طلاب (مقرر ١٠١ إحص) في فصل معين... إلخ.

٢ ـ **غير محدود**: وهو الذي يكون فيه عدد من الأفراد غير منتهٍ (غير محدود) والتي يمكن تمييزها بعضها عن بعض مثل عدد النجوم في سماء يوم صحو، عدد حبات القمح المحصود في مزرعة معينة، عدد طلاب (مقرر ١٠١ إحص) للسنوات العشر القادمة (على فرض استمرار المقرر)... إلخ.

ملاحظة: هناك تقسيم آخر لن نتطرق له في هذا المستوى هو المجتمع الإحصائي غير المحدود وغير المعدود.

في بعض الأحيان يكـون من الصعب ملاحظة بيانات جميع أفراد المجتمع لما يكلف ذلك من جهد ووقت ومال، أو قد يكون في بعض الأحيان استحالة ذلك مثل حصر حبات القمح المحصود، فحص جميع دم المريض وفحص جميع كمية البيض (للتأكد من كونها طازجة). وللتغلب على ذلك يمكن اختيار جزء من المجتمع يسمى بالعيّنة.

وتعرف **العيّنة**: بأنها جزء من المجتمع تختار بحيث تمثل جميع صفات المجتمع. وينفرد بها فرع خاص من علم الإحصاء يسمى نظرية العيّنات وهو خارج نطاق كتابنا هذا. وقد تكون الحاجة ضرورية لأخذ العيّنة بديلاً من دراسة المجتمع كله: مثل أخذ عيّنة من دم شخص لفحصها حيث إننا لا نستطيع فحص كل دم الشخص لأن ذلك يؤدي إلى الوفاة.

وكذلك قد يؤدي دراسة المجتمع كله إلى فقدان عناصره أو إتلافها وهنا يجب أخذ عيّنة صغيرة: فمثلاً عند فحص سلامة كمية من البيض يجب أخذ عينة منها ونقوم بكسرها لنرى ما إن كان البيض سليمًا أم لا. وكذلك عند فحص عمر لمبات لإنتاج مصنع معيّن فإننا نأخذ عيّنة لقياس أعمارها بالإضاءة حتى تحترق، وأفضل العيّنات هي تلك التي تمثل المجتمع أفضل تمثيل، وتفيد المعلومات المتوفرة من العيّنات في التنبؤ عن معلومات ومؤشرات عن المجتمع كله. ومن مميزات العيّنة أنها أقل تكلفة وأكثر سرعة (تستغرق وقتًا أقل) وأكثر شمولاً لإمكانية الحصول على إجابات عن المعلومات المطلوبة بشمول أكبر من الحصر الشامل لأفراد المجتمع محل الدراسة. وكذلك تكون أكثر دقة وذلك بسبب إمكانية استخدام أشخاص ذوي كفاءة عالية ومدربين لأخذ العيّنات من المجتمع محل الدراسة.

(١ ـ ٤) البيانات

هي مجموعة المشاهدات أو الملاحظات المأخوذة أثناء دراسة معيّنة وقد تكون بيانات رقمية (كميّة) مثل أطوال وأوزان مجموعة من الطلاب ودخول مجموعة من الأسر أو بيانات غير رقمية (وصفية) مثل لون البشرة والجنس . . . إلخ .

(١ ـ ٥) المعلمة (Parameter) والإحصائية أو الإحصاءة (Statistic)

المعلمة: شيء يميّز المجتمع كله وذلك مثل متوسط الدخل الشهري للأسر في دولة معيّنة، أو متوسط الطول للذكر البالغ في دولة معيّنة، أو نسبة الذين يدخنون بصفة دائمة في مجتمع معيّن، أو نسبة المعيب في الإنتاج لإحدى السلع وهكذا. . .

أما الإحصائية (الإحصاءة): فهي شيء يميّز العيّنة مثل متوسط الدخل الشهري لعيّنة مكوّنة من ١٠٠ أسرة في دولة ما أو متوسط الطول للذكر البالغ لعيّنة مكونة من ٥٠ ذكرًا وهكذا. . .

(١ ـ ٦) المتغير (Variable)

هو مقدار له خصائص رقمية (كميّة) وغير رقمية (وصفية) تتغير قيمته من عنصر إلى آخر من عناصر المجتمع أو العيّنة. فمثلاً إذا رغبنا في دراسة ظاهرة مثل الوزن أو الطول أو الذكاء أو الجنس أو لون البشرة أو لون الشعر أو لون العيون فإن قراءة المفردات لمتغير الطول أو الوزن أو الذكاء تكون بيانات كميّة أو رقمية (quantitative) وظاهرة الجنس أو لون البشرة أو لون الشعر أو لون العيون تأخذ قيمًا وصفية أو غير رقمية (qualitative) .

(١ ـ ٧) مصادر جمع البيانات الإحصائية

يوجد مصدران لجمع البيانات الإحصائية:

المصدر الأول: تاريخي، وهو ما يؤخذ من السجلات المحفوظة مثل سجلات المواليد والوفيات وإحصاءات هيئة الأمم وغيرها.

المصدر الثاني: ميداني، وهو عبارة عن البيانات المجموعة من أفراد المجتمع كله أو جزء منه بالاتصال المباشر (المقابلة التي يقوم بها العدّادون) أو غير المباشر مثل البريد أو التليفون أو استخدام الطريقين معًا حسب طبيعة المشكلة محل الدراسة.

(١ ـ ٨) الاستمارة الإحصائية

عند دراسة ظاهرة من الظواهر الاقتصادية أو الصناعية أو الزراعية أو الاجتماعية أو الطبية أو غيرها يجب أن تحدد الأهداف لهذه الدراسة حتى يمكن تحديد الأسئلة التي تكون إجابتها كافية للدراسة المطلوبة. توضع هذه الأسئلة في ورقة أو أكثر تسمى بالاستمارة الإحصائية ويراعى عند وضع الأسئلة أن تكون واضحة وملمة بأهداف الدراسة ومختصرة بحيث لا تحذف أي معلومات أساسية وكذلك لا تحتوي على أي تفصيلات غير مطلوبة. وتملأ هذه الاستمارة من مصادر جمع البيانات الإحصائية وهي المصادر التاريخية أو الميدانية التي سبق ذكرها.

وفيما يلي نموذج لاستمارة إحصائية صُمِّمت لغرض المسح الصحي بإحدى مناطق المملكة العربية السعودية:

(١) معلومات شخصية (الجنس) ☐

١ ـ ذكر

٢ ـ أنثى

(٢) العمر (سجل العمر بالسنة) ☐☐

١ ـ أقل من سنة واحدة

٢ ـ من ١ ـ ٣ سنة

٣ ـ من ٤ ـ ٦ سنة

٤ ـ من ٧ ـ ١٢ سنة

٥ ـ من ١٣ ـ ١٨ سنة

٦ ـ من ١٩ ـ ٢٥ سنة

٧ ـ من ٢٦ ـ ٣٤ سنة
٨ ـ من ٣٥ ـ ٤٤ سنة
٩ ـ من ٤٥ ـ ٥٤ سنة
١٠ ـ من ٥٥ ـ ٦٤ سنة
١١ ـ من ٦٥ ـ ٧٤ سنة
١٢ ـ من ٧٥ سنة فما فوق

(٣) الجنسية ☐
١ ـ سعودي
٢ ـ غير سعودي

(٤) الحالة العائلية ☐
١ ـ عــزب
٢ ـ متزوج
٣ ـ مطلق
٤ ـ أرمل
٥ ـ حالات أخرى (حدد)

(٥) المستوى التعليمي ☐
١ ـ أمــي
٢ ـ متعلم بدون شهادة
٣ ـ أنهى الدراسة الابتدائية
٤ ـ أنهى الدراسة المتوسطة
٥ ـ أنهى الدراسة الثانوية
٦ ـ أنهى الدراسة الفنية
٧ ـ أنهى الدراسة الجامعية

(٦) **مكان الولادة** **٧ ـ ٩** ☐☐☐

١ ـ الأحســـاء
٢ ـ المنطقة الشرقية
٣ ـ المنطقة الوسطى
٤ ـ المنطقة الغربية
٥ ـ جـــيزان
٦ ـ القصـــيم
٧ ـ المنطقة الشمالية
٨ ـ البحرين
٩ ـ قـطـــر
١٠ ـ الإمـارات
١١ ـ اليمـــن
١٢ ـ العـــراق
١٣ ـ مـصـــر
١٤ ـ الســودان
١٥ ـ بلاد عربية أخرى
١٦ ـ إيـــران
١٧ ـ آسـيـــا
١٨ ـ أفريقيا
١٩ ـ المملكة المتحدة
٢٠ ـ أوروبـا
٢١ ـ الولايات المتحدة الأمريكية
٢٢ ـ المكسيـك
٢٣ ـ جنوب أمريكا
٢٤ ـ كـنـــدا

٢٥ ـ أستراليا ـ نيوزيلندا
٢٦ ـ أي بلد آخر

(٧) الغرض من السفر خارج القرية أو المدينة : **١٠ ـ ١٢** ☐☐☐

١ ـ الذهاب إلى مكان العمل
٢ ـ ذهابًا إلى المدرسة
٣ ـ الذهاب للعمرة أو الحج
٤ ـ ذهابًا إلى المطار
٥ ـ ذهابًا إلى الطبيب
٦ ـ عائدًا إلى البيت
٧ ـ زيارة اجتماعية
٨ ـ للاستجمام
٩ ـ أشياء أخرى (حـدد)

(٨) هل سافرت إلى المنطقة الجنوبية من المملكة؟ **١٣** ☐

١ ـ نعـم
٢ ـ لا

(٩) إن كانت الإجابة نعم فمتى تمت الزيارة؟ **١٤** ☐

١ ـ خلال الـ ١٥ يومًا السابقة
٢ ـ خلال الشهر السابق

(١٠) عـادات النـوم **١٥** ☐

في الصيف يكون معظم النوم في :
١ ـ داخل المنزل
٢ ـ خارج المنزل
٣ ـ الاثنين معًا بالقدر نفسه .

(١١) عند النوم في الصيف داخل المنزل فأين تنام؟ ١٦ ☐

١ - في غرفة النوم في المنزل

٢ - في غرفة بها حيوان

٣ - في حجرة في المزرعة

(١٢) إن كان النوم في الخارج صيفًا فكيف تنام؟ ١٧ ☐

١ - يكون النوم على الأرض من غير فراش

٢ - يكون النوم على الأرض بفراش

٣ - يكون النوم أحيانًا على الأرض من غير فراش وأحيانًا على فراش

(١٣) إن كان النوم في الخارج صيفًا ١٨ ☐

١ - يستعمل مروحة دائمًا

٢ - يستعمل مروحة أحيانًا

٣ - لا يستعمل مروحة

(١٤) آخر الأمراض التي أصيب بها الفرد ١٩ ☐

١ - لا

٢ - نعم مرة

٣ - نعم مرتين

٤ - نعم ثلاث مرات

٥ - نعم أربع مرات

٦ - نعم أكثر من أربع مرات

٧ - لا أعرف عدد المرات

٨ - لا أعرف قط

(١٥) أسباب المرض ٢٠ - ٢١ ☐☐

١ - برد أو التهاب بالحلق

٢ ـ مرض بالأذن
٣ ـ مرض بالعيون
٤ ـ مرض بالأسنان
٥ ـ ربـــو
٦ ـ التهابات أخرى بالصدر
٧ ـ حـمـى
٨ ـ نزلة معدية أو معوية (إسهال وألم بالبطن وقيء)
٩ ـ آلام بالمفاصل
١٠ ـ حادث أو جرح
١١ ـ حساسية بالجلد
١٢ ـ حـب الشباب
١٣ ـ أمراض جلدية أخرى
١٤ ـ أمراض أخرى (حـدد)

(١٦) أنواع الخدمات الطبية التي لجأ إليها الفرد عندما كان مريضًا

١ ـ طبيب عام ٢٢ ☐
٢ ـ مستوصف حكومي
٣ ـ مستشفى حكومي
٤ ـ مستشفى خصوصي
٥ ـ كل ما سبق (١+٢+٣+٤)
٦ ـ لا شيء مما سبق ويستعمل علاجًا بالمنزل أو من صديق

(١٧) الصعوبات في الاستفادة من الخدمات الطبية ٢٣ ☐

١ ـ بعيدة جدًا
٢ ـ تأخذ وقتًا كثيرًا
٣ ـ صعوبة المواصلات

٤ ـ أمور مالية

٥ ـ غير مقتنع بما يقدم من خدمات

٦ ـ صعوبات أخرى (حــدد)

(١٨) من يقوم بتغطية تكاليف الخدمات الطبية؟ ٢٤ ☐

١ ـ الدولة

٢ ـ شركـة

٣ ـ بنفســه

٤ ـ جهات أخرى (حدد)

(١ ـ ٩) تمــارين

(١) عرِّف علم الإحصاء؟.

(٢) أذكر أمثلة عن المجتمع الإحصائي والعيّنة الإحصائية.

(٣) ما هي مصادر جمع المعلومات في المنطقة التي تُقِيم فيها مع الأمثلة؟

(٤) أذكر ما تعرفه عن: المعلمة ـ الإحصائية (الإحصاءة) ـ المتغير ـ البيانات. ثم مثِّل لكل منها.

(٥) ضع العلامة (√) أمام كل من العبارات الآتية إذا كانت صحيحة أو العلامة (X) إذا كانت خاطئة:

ا ـ العينة الإحصائية هي جزء من المجتمع الإحصائي ☐

ب ـ المعلمة تمثل خاصة من خواص المجتمع ☐

جـ ـ العينة هي دائمًا هدف الدراسة والمجتمع هو الذي يخضع بالفعل للدراسة ☐

(٦) أكمل ما يلي:

ا ـ الاستمارة الإحصائية هي: ....

ب ـ مصدر جمع المعلومات التاريخي هو: ....

جـ ـ مصدر جمع المعلومات الميداني هو: ....

د ـ الخــاصــية التي تـصـف مجتمعًــا تسمى ....

وتسمى .... إذا كانت تصف عينة.

٢

تنظيم البيانات وعرضها

● تنظيم البيانات وتلخيصها وعرضها جدوليًا
● العرض البياني ● الرسوم البيانية ● تمارين

(٢ - ١) تنظيم البيانات وتلخيصها وعرضها جدوليًا

بعد جمع البيانات سواء من المصادر التاريخية أو المصادر الميدانية فإنها تكون بيانات أولية غير منتظمة عدديًا وتصعب دراستها أو استنتاج أي شيء منها. ولذلك دعت الحاجة إلى تنظيم وتلخيص هذه البيانات بصورة يسهل فهمها واستنتاج بعض النتائج الأولية منها ولتوضيح ذلك نعتبر المثال الآتي:

مثال (٢ - ١)

إذا كان لدينا تقديرات ٦٠ طالبًا كالتالي:

D	B	E	C	D	B	D	C	E	A
B	E	C	D	B	D	D	A	E	C
C	D	A	C	E	D	C	C	D	B
D	E	D	D	A	D	D	C	D	C
D	A	B	D	B	D	C	D	C	E
D	B	C	C	E	D	C	C	D	A

والبيانات السابقة بوضعها الحالي قد تجعل من الصعب التعرف على الطلاب الحاصلين على تقدير مشترك مثل ممتاز (A) أو جيدجدًا (B) . . . . ومن هنا أصبحت الحاجة إلى وضع التقديرات وتلخيصها في جدول يسهل دراسته يسمى بجدول التوزيع التكراري وقد تكون البيانات رقمية مثل درجات الطلاب، أو أطوال الطلاب، أو أوزانهم، أو أجور العمل في أحد المصانع. ونوضح ذلك بالمثال التالي:

مثال (٢ ـ ٢)

البيانات التالية تمثل درجات ٥٠ طالبًا في إحدى المواد:

51	95	70	74	73	90	71	74	90	67
91	72	83	89	50	80	72	84	85	69
62	82	87	76	91	76	87	75	78	79
71	96	81	88	64	82	73	57	86	70
80	81	75	85	74	90	83	66	77	91

البيانات السابقة بوضعها الحالي يصعب دراستها أو استنتاج بعض المؤشرات منها. فمثلاً ما هو عدد الطلبة الذين حصلوا على 70 درجة فأكثر؟ أو عدد الطلبة الذين حصلوا على درجات تتراوح ما بين 80 درجة و 90 درجة .... إلخ؟ ولذلك فإن أول مرحلة للتحليل الإحصائي تتكون من تصميم جدول التوزيع التكراري، وقبل التعرض لكيفية تنظيم هذه البيانات في جداول تكرارية يلزم أن نعرف أن البيانات الإحصائية نوعان: بيانات وصفية (qualitative data) وبيانات كمية (quantitative data).

(٢ ـ ١ ـ ١) البيانات الوصفيّة (qualitative data)

وهي البيانات التي تصف الأفراد والمجتمع مثل: لون الشعر أو العيون أو البشرة أو تقديرات النجاح للطلاب في إحدى المواد كما ورد في مثال (٢ ـ ١) السابق.

(٢ ـ ١ ـ ٢) البيانات الكميّة (quantitative data)

وهي البيانات التي يقاس فيها الأفراد والمجتمع بمقاييس كمية (رقمية) مثل: أطوال الطلاب وتقاس بالسنتيمتر، وأوزان الطلاب تقاس بالكيلوجرام، وأعمار الطلاب تقاس بالسنة، أما نتيجة الامتحان فبالدرجات وأجور العمال بالريال.

وتنظم وتلخص البيانات الإحصائية سواء أكانت وصفية أم كمية فيما يسمى بالتوزيع (الجدول) التكراري (frequency distribution) وهو عبارة عن جدول يلخص البيانات الأولية فيوزعها على فئات أو طوائف ويحدد عدد الأفراد الذين ينتمون إلى كل فئة ويسمى هذا العدد تكرار الفئة؛ وعادة يرمز له بـ f ولإتمام ذلك ينبغي أن يُعمل

جدول آخر يسمى جدول تفريغ البيانات الإحصائية وهو يتكون من ثلاث خانات. الخـانة الأولى أو العمود الأول تكتب فيه الصفة للبيانات الوصفية أو الفئة للبيانات الكميّة وفي الخـانة الثانية توضع العلامات وهي عبارة عن حِزَمٍ كل حزمة مكونة من خمسة خطوط، أربعة منها رأسية والخامس مائل يحزم الخطوط الأربعة الرأسية وبذلك تصبح الحزمة على الصورة 𝍸 وفي الخانة الثالثة الأخيرة يكتب مجموع العلامات أمام كل صفة أو فئة على حدة ومجموع هذه العلامات في كل فئة تسمى التكرار لهذه الصفة أو الفئة. وبذلك يكون جدول تفريغ البيانات الإحصائية الوصفية في مثال (٢ ـ ١) وهو تقديرات النجاح للطلاب في إحدى المواد كالتالي:

جدول رقم (٢ ـ ١) يبين تفريغ وتوزيع التقديرات للطلاب في مثال (٢ ـ ١)

الصفات	العلامات	التكرار (عدد الطلاب)
A	𝍸 I	6
B	𝍸 III	8
C	𝍸 𝍸 𝍸 I	16
D	𝍸 𝍸 𝍸 𝍸 II	22
E	𝍸 III	8
المجموع		60

ومن هذا الجـدول نكـوّن جدولاً آخـر يسمى الجدول التكراري أو جدول التوزيع التكراري للبيانات الوصفية الذي يتكون من خانتين. الأولى تمثل الصفة والثانية تمثل التكرار كما هو مبين بجدول رقم (٢ ـ ٢) كما يلي:

جدول رقم (٢-٢) يبين التوزيع التكراري
لتقديرات الطلاب في مثال (٢-١)

الصفـات	التكـرار
A	6
B	8
C	16
D	22
E	8
المجموع	60

وأحيانًا يكتب الجدول السابق رقم (٢ ـ ٢) في صورة أفقية كما يلي:

جدول رقم (٢ ـ ٣) يبين التوزيع التكراري لتقديرات الطلاب في مثال (٢ ـ ١)

الصفـة	A	B	C	D	E	المجموع
التكرار	6	8	16	22	8	60

وبعد إلقاء الضوء على كيفية عمل التكرارات أمام الصفات وتكوين الجداول التكرارية للبيانـات الـوصفية في الجـداول السـابقـة فإنه يلزم قبل الدخول في عمل الجداول التكرارية للبيانات الكميّة شرح كيفية تكوين الفئات أو الفترات المنتظمة (المتساوية الطول) كما يلي.

(٢ ـ ١ ـ ٣) طريقة عمل الفئات (classes) المنتظمة للبيانات الكميّة

الغرض من عمل الفئات هو تجميع القيم المتقاربة في مجموعات ولا توجد هناك قواعد ثابتة لتحديد طول الفئات وعددها إلا إنه من المرغوب فيه ألا يكون عدد الفئات

صغيراً فتضيع معالم التوزيع وتفقد كثيراً من التفاصيل. كما لا يكون عدد الفئات كبيراً جدًّا فتضيع الحكمة من التجميع في فئات، وعادة يتراوح عدد الفئات من 5 إلى 20 فئة. ولتحديد عدد الفئات وطول كل فئة فإنه يعتمد إلى حد كبير على الخبرة ومدى البيانات (range) وهو الفرق بين أكبر قراءة وأصغر قراءة كحد أقصى ولتوضيح كيفية عمل الفئات المنتظمة نعتبر مثال (٢ ـ ٢) السابق وتكون الخطوات كالتالي:

أ ـ نحسب طول المدى للقراءات R أي أن:

$$R = 96 - 50 = 46$$

ب ـ نختار مثلاً عدد الفئات = 5 فئات.

جـ ـ نحسب طول الفئة بأن نقسم المدى على عدد الفئات 5 بحيث يقرّب الكسر إن وجد من خارج القسمة إلى الواحد الصحيح في مثالنا هذا مهما كانت قيمة الكسر وبذلك يكون طول الفئة L عددًا صحيحًا أي أن:

$$L = 46 \div 5 = 9.4$$

ويقرّب فيصبح طول الفئة $L = 10$. إذا كانت البيانات تحتوي على مشاهدات كسرية يقرّب طول الفئة للرقم العشري المعطاة به البيانات.

د ـ تُختار أصغر قراءة في البيانات لتكون بداية الفئة الأولى المقرّبة ويضاف إليها طول الفئة فنحصل على بداية الفئة الثانية وفي المثال (٢ ـ ٢) بداية الفئة الأولى المقرّبة 50 فتكون بداية الفئة الثانية هي:

$$50 + 10 = 60$$

هـ ـ تحدد بداية الفئة الثالثة المقرّبة بإضافة طول الفئة لبداية الفئة الثانية المقرّبة وهكذا لباقي الفئات.

و ـ لإيجاد نهاية أي فئة نضيف إلى بدايتها طول الفئة مطروحًا منه واحد (وحدة الدقة في مثالنا)، وفي هذا المثال تكون نهاية الفئة الأولى المقرّبة هي 59 ، ونهاية الفئة الثانية المقرّبة 69 ، وهكذا لباقي الفئات، ويكون جدول تفريغ البيانات كما هو موضح بالجدول التالي:

جدول رقم (٢ ـ ٤) يبين تفريغ وتوزيع الدرجات للطلاب في مثال (٢ ـ ٢)

التكرار (عدد الطلاب)	العلامات	الصفات
3	III	50 - 59
5	𝍸	60 - 69
18	𝍸 𝍸 𝍸 III	70 - 79
16	𝍸 𝍸 𝍸 I	80 - 89
8	𝍸 III	90 - 99
50		المجموع

ويلخص من جدول التفريغ رقم (٢ ـ ٤) جدول التوزيع التكراري للبيانات الإحصائية الكميّة الذي يتكون من خانتين. الأولى يكتب بها حدود الفئات والثانية التكرار كما هو مبين في الجدول الآتي:

جدول رقم (٢ ـ ٥) يبين التوزيع التكراري
لدرجات الطلاب في مثال (٢ ـ ٢)

التكرار	حدود الفئات
3	50 - 59
5	60 - 69
18	70 - 79
16	80 - 89
8	90 - 99
50	المجموع

والجدول السابق رقم (٢ ـ ٥) يمكن أن يكتب في صورة أفقية لتوفير حيز الكتابة كالآتي:

المجموع	90-99	80-89	70-79	60-69	50-59	الفئات
50	8	16	18	5	3	التكرار

ويمكن تكوين جدولين آخرين من جدول التوزيع التكراري رقم (٢ ـ ٥) وهما الجدول التكراري النسبي (relative frequency table) والجدول التكراري المئوي (percentage frequency table) . فالجدول التكراري النسبي للبيانات الإحصائية يتكون من خانتين مثل الجدول التكراري العادي ولكن خانة التكرار يكتب بها التكرار النسبي : وهو عبارة عن التكرار لأي فئة مقسومًا على مجموع التكرارات ويكون مجموع التكرارات النسبية لجميع الفئات يساوي واحدًا صحيحًا ويوضح الجدول من مثال (٢ ـ ٢) كالآتي:

جدول رقم (٢ ـ ٦) يبين التوزيع التكراري النسبي لدرجات الطلاب في المثال (٢ ـ ٢)

التكرار النسبي	حدود الفئات
0.06	50 – 59
0.10	60 – 69
0.36	70 – 79
0.32	80 – 89
0.16	90 – 99
1	المجموع

والجدول التكراري المئوي للبيانات الإحصائية يتكون من خانتين أيضًا مثل الجدول التكراري النسبي السابق ولكن في خانة التكرارات النسبية تكتب التكرارات المئوية

ويمكن الحصول عليها بضرب التكرار النسبي في 100 ونلاحظ أن مجموع التكرارات المئوية يساوي 100 وبذلك يكون الجدول التكراري المئوي للبيانات في مثال (٢ ـ ٢) كالتالي:

جدول رقم (٢ ـ ٧) يبين التوزيع التكراري المئوي لدرجات الطلاب في مثال (٢ ـ ٢)

التكرار المئوي	حدود الفئات
6	50 – 59
10	60 – 69
36	70 – 79
32	80 – 89
16	90 – 99
100	المجموع

(٢ ـ ١ ـ ٤) الحدود الحقيقية (الفعلية) للفئات

البيانـات الإحصائية المراد تلخيصها وتنظيمها في جداول تكرارية عادة تكون مكتوبة مقرّبة مثلاً لأقرب وحدة قياس أو لأقرب نصف وحدة قياس فإذا كانت البيانات مقرّبة لأرقام صحيحة فإننا نطرح من الحد الأدنى المقرّب للفئة 0.5 لنحصل على الحد الأدنى الحقيقي ونضيف 0.5 إلى الحد الأعلى المقرب لنحصل على الحد الأعلى الحقيقي للفئة وهكذا لباقي الفئات للحصول على الحدود الحقيقية لها. أما إذا كانت البيانات محسوبة لأقرب رقم عشري فإننا نطرح 0.05 من الحد الأدنى المقرّب للفئة لنحصل على الحـد الأدنى الحقيقي لها ونضيف 0.05 إلى حدهـا الأعلى المقرّب لنحصل على الحد الأعلى الحقيقي لهذه الفئة وهكذا لباقي الفئات. وبالمثل يمكن إيجاد أي حدود حقيقية مهما كانت أعداد الأرقام العشرية المقرّبة بالطريقة السابقة نفسها وبذلك يكون جدول التوزيع التكراري رقم (٢ ـ ٥) مستخدمًا الحدود الحقيقية للفئات كالآتي في جدول رقم (٢ ـ ٨).

جدول رقم (٢ ـ ٨) يبين التوزيع التكراري لدرجات الطلاب بالحدود الفعلية (الحقيقية) للفئات في مثال (٢ ـ ٢)

التكرار	الحدود الحقيقية للفئات
3	49.5–59.5
5	59.5–69.5
18	69.5–79.5
16	79.5–89.5
8	89.5–99.5
50	المجمـــوع

(٢ ـ ١ ـ ٥) مركز الفئات (class mark)

يعرف مركز الفئة بالعلاقة الآتية:

$$\text{مركز الفئة} = \frac{\text{الحد الأعلى للفئة + الحد الأدنى للفئة نفسها}}{2}$$

ومركز الفئة لا يتأثر بحدود الفئات سواء أكانت حدودًا مقرّبة أو حدودًا حقيقية ففي مثـال (٢ ـ ٢) جدول (٢ ـ ٥) نحسب منـه مركـز الفئـة الأولى من الحـدود المقرّبة $\frac{50+59}{2}=54.5$ درجــة ، ومن جدول (٢ ـ ٨) نحسب مركـز الفئـة من الحـدود الحقيقية $\frac{59.5\ +49.5}{2}=54.5$ درجـة ويتبـع الـطريقـة نفسها لحساب باقي مراكز الفئات الأخرى أو بإضافة طول الفئة إلى مركز الفئة الأولى لنحصل على مركز الفئة الثانية وهكذا لباقي الفئات ويمكن تنظيم الجداول السابقة للبيانات الكمية في جدول واحد يشمل الحدود المقربة والحقيقية للفئات ومراكز الفئات والتكرار والتكرار النسبي والتكرار المئوي ومن مثال (٢ ـ ٢) يكون جدول شامل لتوزيع درجات الطلاب كالتالي:

جدول رقم (٢ - ٩) يبين التوزيع التكراري لدرجات الطلاب في مثال (٢ - ٢) وتلخيص للجداول السابقة

التكرار المئوي	التكرار النسبي	التكرار	مراكز الفئات	الحدود الحقيقية	الحدود المقرّبة للفئات
6	0.06	3	54.5	49.5 – 59.5	50 – 59
10	0.10	5	64.5	59.5 – 69.5	60 – 69
36	0.36	18	74.5	69.5 – 79.5	70 – 79
32	0.32	16	84.5	79.5 – 89.5	80 – 89
16	0.16	8	94.5	89.5 – 99.5	90 – 99
100	1.00	50			المجموع

ملاحظة مهمة :

لسهولة بناء الجداول الإحصائية المستنتجة من الجدول التكراري وكذلك الحسابات الإحصائية المختلفة التي سوف نتعرض لها بالشرح فيما بعد يجب أن تكون حدود الفئات في الجداول التكرارية حدودًا حقيقية .

(٢ - ١ - ٦) الجدول التكراري المتجمع الصاعد (''less than'' cumulative frequency)

في كثير من الأحيان يكون اهتمامنا منصبًا على عدد القراءات التي تكون أصغر من أو تساوي مقدارًا معيّنًا ، ففي مثال (٢ - ٢) يمكن أن يطلب ما هو عدد الطلاب الحاصلين على 79 درجة فأقل؟ فتكون الإجابة :

عدد الطلاب الحاصلين على 79 درجة فأقل = 3 + 5 + 18 = 26 طالبًا .

وهذا هو التكرار المتجمع الصاعد للفئة الثالثة . وكذلك يمكن استخدام الجدول في إيجاد عدد الطلاب الذين تنحصر درجاتهم بين حدين معلومين . ويمكن كتابة الجدول التكراري المتجمع الصاعد المكون من خانتين . **الأولى** : يكتب في السطر الأول منها أقل من الحد الأدنى الحقيقي للفئة الأولى (بدلاً من حدود الفئة الأولى) وكذلك لباقي الفئات حتى نصل إلى الفئة الأخيرة فيكتب لها سطرين الأول منهما: أقل من الحد

الأدنى الحقيقي للفئة الأخيرة والثاني منها أقل من الحد الأعلى الحقيقي للفئة الأخيرة كما سيوضح في جدول (٢ ـ ١٠) للتوزيع التكراري المتجمع الصاعد لدرجات الطلاب في مثال (٢ ـ ٢).

جدول رقم (٢ ـ ١٠) يبين التوزيع التكراري المتجمع الصاعد لدرجات الطلاب في مثال (٢ ـ ٢)

التكرار المتجمع الصاعد	حدود الفئات	
0	أقل من	49.5
3	أقل من	59.5
8	أقل من	69.5
26	أقل من	79.5
42	أقل من	89.5
50	أقل من	99.5

(٢ ـ ١ ـ ٧) الجدول التكراري المتجمع الهابط ("or more" cumulative frequency)

قد يكون اهتمامنا أحيانًا منصبًّا على عدد القيم التي تكون أكبر من أو تساوي قيمة معيّنة ففي مثال (٢ ـ ٢) قد يطلب معرفة عدد الطلاب الحاصلين على 79 درجة فأكثر فتكون الإجابة هي :

عدد الطلاب الذين حصلوا على 79 درجة فأكثر = 8 + 16 = 24 طالبًا

والجدول التكراري المتجمع الهابط مثل الجدول التكراري المتجمع الصاعد مكون من خانتين الأولى يكتب في السطر الأول منها أكبر من الحد الأدنى الحقيقي للفئة الأولى (بدلاً من حدود الفئة الأولى) وهكذا لباقي الفئات حتى نصل إلى الفئة الأخيرة فيكتب سطران آخران أولهما أكبر من الحد الأدنى الحقيقي للفئة الأخيرة وثانيهما يكتب فيه أكبر من الحد الأعلى الحقيقي للفئة الأخيرة وجدول رقم (٢ ـ ١١) يوضح التوزيع التكراري المتجمع الهابط لدرجات الطلاب في مثال (٢ ـ ٢).

جدول رقم (٢ ـ ١١) يبين التوزيع التكراري الهابط لدرجات الطلاب في مثال (٢ ـ ٢)

التكرار المتجمع الهابط	حدود الفئات
50	أكبر من 49.5
47	أكبر من 59.5
42	أكبر من 69.5
24	أكبر من 79.5
8	أكبر من 89.5
0	أكبر من 99.5

(٢ ـ ١ ـ ٨) الفئات غير المنتظمة

سبق الكلام عن الفئات المنتظمة وهي غالبًا ما تكون ذات أهمية كبيرة وخاصة في العمليات الإحصائية التي سوف نتعرض لها فيما بعد وذلك لسهولتها في التطبيق بدون تعديل التكرار لها. ولكننا أحيانًا نضطر إلى استخدام فئات غير منتظمة في بعض الظواهر محل الدراسة لأن الفئات المنتظمة قد لا تفي بالغرض وذلك بأن يكون تكرارها قليلاً أو خاليًا من التكرار مثل ظاهرة الدخول للأفراد أو الأجور أو درجات الامتحان للطلاب أو الوفيات للأطفال الرضع (أقل من سنة). فإن عمل جداول ذات فئات غير منتظمة يكون مناسبًا. ولكن عند رسم المدرج التكراري وغيره من الرسوم البيانية فإنه يتطلب تعديل التكرار للفئات غير المنتظمة حتى يصبح الرسم ممثلاً لهذه البيانات. وسوف نتناول طريقة تعديل التكرارات عند رسم المدرج التكراري فيما بعد.

(٢ ـ ١ ـ ٩) الجداول التكرارية الثنائية أو المزدوجة (bivariate frequency tables)

في بعض الأحيان تكـون البيانـات لأكثـر من متغـير للوحدات محل الدراسة الإحصائية. فإذا كان لدينا مجموعة من الطلاّب ونرغب في دراسة ظاهرة الطول وظاهرة الوزن فيهم أو دراسة درجات اختبارين لمادتين مختلفتين لهم أيضًا. أو دراسة الأجور والإنتاج لمجموعة من العمّال في إحدى المؤسسات. ففي مثل هذه الحالات فإنه يلزم منا

عمل جداول توزيع تكرارية مزدوجة تظهر فيها تكرار كل من الظاهرتين محل الدراسة. وفي الجداول التكرارية المزدوجة تكتب حدود الفئات في وضع رأسي للظاهرة الأولى وحدود الفئات للظاهرة الثانية في وضع أفقي. ويكون الجدول المزدوج عبارة عن شبكة من المربعات أو مصفوفة (matrix) في صورة صفوف أفقية وأعمدة رأسية ويكتب التكرار المشترك للظاهرتين داخل هذه المربعات بحيث يكون بداية الصف هو الحد الأدنى للفئة الظاهرة الأولى وبداية العمود هو الحد الأدنى للفئة الظاهرة الثانية وفي نهاية كل من الصف والعمود يكتب مجموع التكرار لكل من الصف والعمود وبذلك تكون التكرارات الرأسية في خانة المجموع تمثل تكرارات الظاهرة الأولى والتكرارات الأفقية في خانة المجموع تمثل التكرارات للظاهرة الثانية ونوضح ذلك بالمثال التالي.

مثال (٢ - ٣)

الجدول الآتي يمثل درجات 30 طالبًا في كل من مادتي الإحصاء والرياضيات والمطلوب عمل جدول توزيع تكراري لهذه البيانات.

إحصاء	رياضيات	إحصاء	رياضيات	إحصاء	رياضيات	إحصاء	رياضيات	إحصاء	رياضيات
50	55	80	75	53	50	57	90	76	71
70	72	71	68	72	65	73	75	93	93
81	80	62	65	85	86	90	92	64	67
61	60	83	82	56	52	74	72	94	96
82	85	63	60	86	81	91	92	77	72
79	75	84	81	60	57	75	70	78	77

الحـــل

ننشىء جدولاً للتفريغ مزدوجًا بحيث نختار أطوالاً مناسبة لحدود الفئات في كل من الإحصاء والرياضيات وفي هذا المثال تساوى عشر درجات وتكتب فئات درجات الإحصـاء رأسيًا وفئات درجات الرياضيات أفقيًا وتفرغ الدرجات بالعلامات فمثلاً الطالب الحاصل على 71 درجة في الإحصاء و 68 درجة في الرياضيات توضع له علامة

في المـربـع الذي يبدأ بحدود الفئات في الإحصاء (79 – 70) للصف وحدود الفئات للرياضيات (69 – 60) للعمود وتكرر هذه العملية لباقي الطلاب فنحصل على جدول تفريغ البيانات المزدوجة كما هو موضح في الجدول التالي.

جدول رقم (٢ ـ ١٢) يبين تفريغ البيانات المزدوج لمادتي الإحصاء والرياضيات لمثال (٢ ـ ٣)

الرياضيات / الإحصاء	50 – 59	60 – 69	70 – 79	80 – 89	90 – 99	المجموع
50 – 59	III					3
60 – 69	I	𝍸				5
70 – 79		II	𝍸 III			10
80 – 89			I	𝍸 I	I	8
90 – 99					𝍸	4
المجموع	4	6	9	6	5	30

وبعد الانتهاء من جدول التفريغ المزدوج يصاغ الجدول التكراري المزدوج منه وذلك بأن يوضع بدلاً من العلامات مقدار ما تساويه بالأرقام، ويفيد هذا الجدول في إظهار العـلاقـة بين الـظاهرتين من الترابط وعدم الترابط كما يتضح ذلك فيما بعد في باب الارتباط والانحدار.

ويصبح جدول التوزيع التكراري المزدوج كالتالي:

جدول رقم (٢ ـ ١٣) يبين التوزيع التكراري المزدوج لدرجات الطلاب في الإحصاء والرياضيات في مثال (٢ ـ ٣)

الرياضيات / الإحصاء	50-59	60-69	70-79	80-89	90-99	المجموع
50-59	3					3
60-69	1	4				5
70-79		2	8			10
80-89			1	6	1	8
90-99					4	4
المجموع	4	6	9	6	5	30

وقد تكون البيانات الثنائية وصفية لتقديرات مجموعة من الطلاب في مادتين مختلفتين مثل الكيمياء والرياضيات فإننا نتّبع الخطوات نفسها التي تمت في الجداول المزدوجة للبيانات الكميّة وتوضح ذلك بالمثال الآتي.

مثال (٢ ـ ٤)

البيانات الآتية تمثل تقديرًا لـ 20 طالبًا في مادتي الكيمياء والرياضيات. أعرض هذه البيانات في شكل جدول تكراري مزدوج.

الكيمياء	C	C	D	E	A	A	B	C	C	B
الرياضيات	C	B	B	E	A	C	C	B	C	C
الكيمياء	B	C	C	A	D	A	C	B	B	A
الرياضيات	B	A	B	B	E	A	A	C	A	C

الحـــل

ننشىء جدول تفريغ البيانات (٢ ـ ١٤) التالي كما سبق في مثال (٢ ـ ٣)

جدول رقم (٢ ـ ١٤) تفريغ البيانات المزدوج للبيانات الوصفية مثال (٢ ـ ٤)

الرياضيات / الكيمياء	A	B	C	D	E	المجموع
A	II	I	III			6
B	I	I	III	I		6
C	I	III	II			6
D						0
E				I	I	2
المجموع	4	5	8	2	1	20

ويصاغ من الجدول السابق جدول التوزيع التكراري المزدوج رقم (٢ ـ ١٥) كالتالي:

جدول رقم (٢ ـ ١٥) يبين التوزيع التكراري المزدوج للبيانات في مثال (٢ ـ ٤)

الرياضيات / الكيمياء	A	B	C	D	E	المجموع
A	2	1	3			6
B	1	1	3	1		6
C	1	3	2			6
D						0
E				1	1	2
المجموع	4	5	8	2	1	20

(٢ ـ ٢) العرض البياني (Graphical representation)

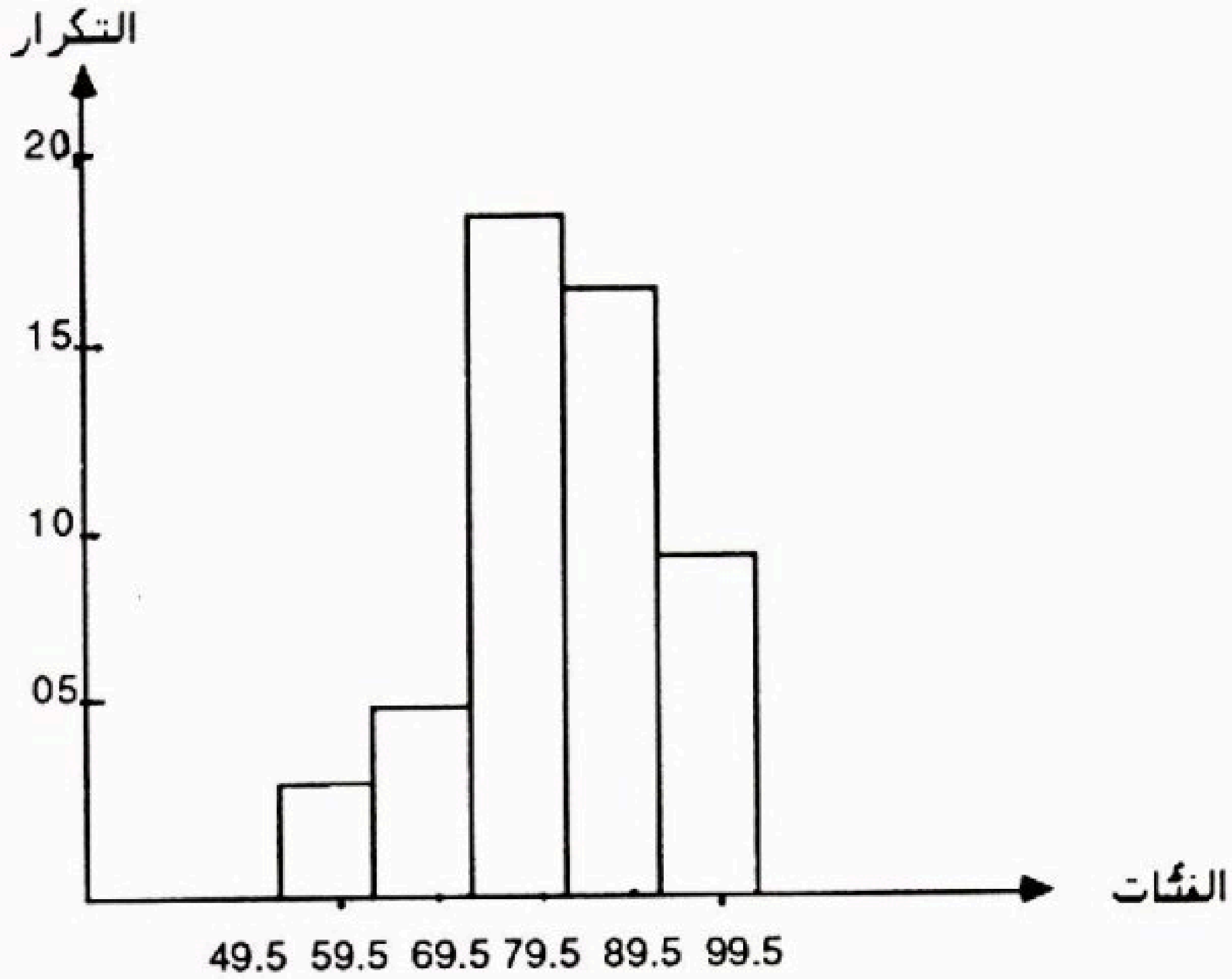
لقد تكلمنا عن طرق تنظيم وتلخيص البيانات وعرضها جدوليًا وقد لاحظنا أن عرض البيانات في صورة جداول تكرارية يعطي صورة شاملة واضحة عن البيانات الأولية وتوزيعاتها التكرارية. ومع ذلك فإن عرض الجداول التكرارية بالتمثيل البياني يعطي فكرة أوضح وأسرع عن أشكال التوزيعات التكرارية وبذلك يمكن عرض التوزيعات التكرارية بيانيًا باستخدام:

- المدرج التكراري (histogram)
- المضلع التكراري (polygon)
- المنحنى التكراري (frequency curve)
- المنحنى التكراري المتجمع الصاعد والهابط (cumulative frequency curve)

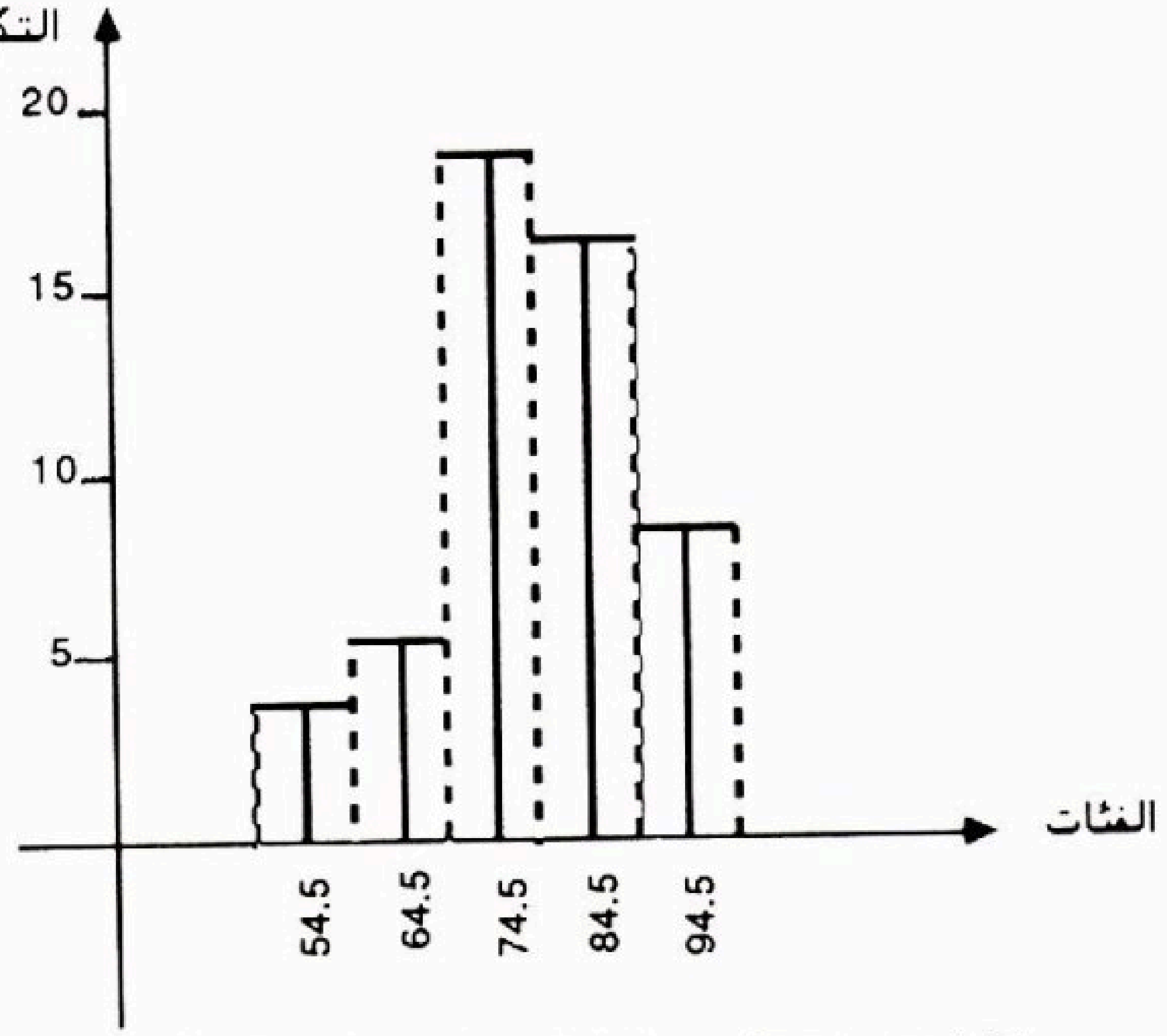
(٢ ـ ٢ ـ ١) المدرج التكراري (histogram)

نرسم المدرج التكراري على محورين متعامدين أحدهما أفقي يمثل الفئات والثاني رأسي يمثل التكرار. ونرسم مستطيلات متلاصقة على الفئات قاعدتها طول الفئة محسوبًا من الحدود الحقيقية. وارتفاعاتها عبارة عن تكرار هذه الفئات. فمثلاً بالنسبة للفئة الأولى يكون المستطيل قاعدته بادئة من الحد الأدنى الحقيقي للفئة الأولى، ومنتهية في الحد الأعلى للفئة الأولى. وارتفاع المستطيل هو تكرار الفئة الأولى. وهكذا لباقي المستطيلات التي تمثل باقي التكرارات والمدرج التكراري للبيانات الموجودة في الجدول (٢ ـ ٥) يوضح بشكل (٢ ـ ١).

ويمكن رسم المدرج التكراري بطريقة أخرى. وهي أن تحدد مراكز الفئات على المحور الأفقي ومنها يرسم ارتفاع المستطيل الممثل للتكرار في منتصف القاعدة للمستطيل على أن يكون البعد من أحد جوانب مركز الفئة مساويًا لبعد الجانب الآخر ويستكمل رسم المستطيل للفئة الأولى ويتبع نفس الطريقة لباقي الفئات ويوضح رسم المدرج التكراري لجدول (٢ ـ ٥) بهذه الطريقة كما هو مبين بشكل (٢ ـ ٢).



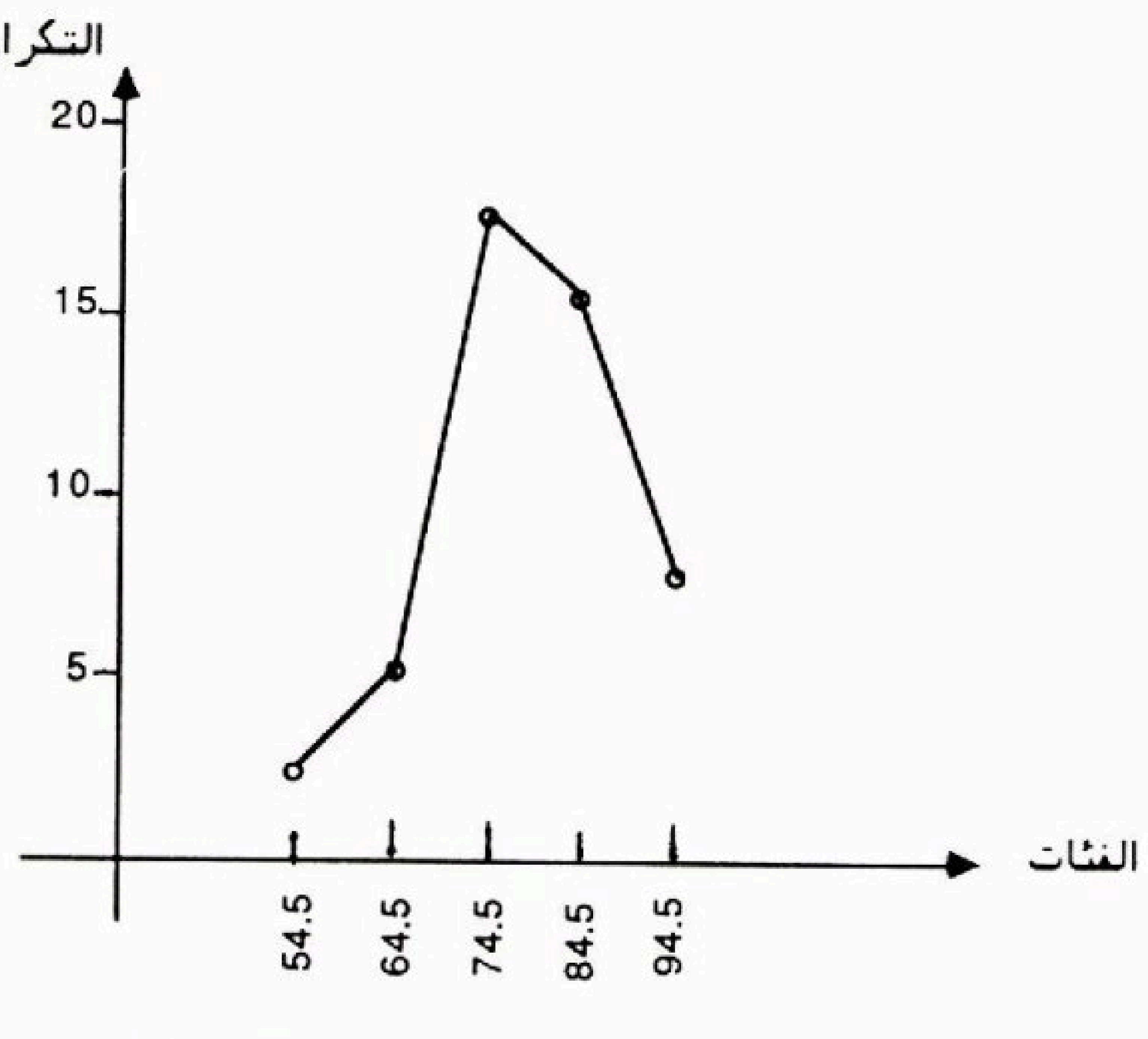
شكل (٢ - ١) المدرج التكراري باستخدام الحدود الفعلية للفئات



شكل (٢ - ٢) المدرج التكراري باستخدام مراكز الفئات

(٢ - ٢ - ٢) المضلع التكراري (polygon)

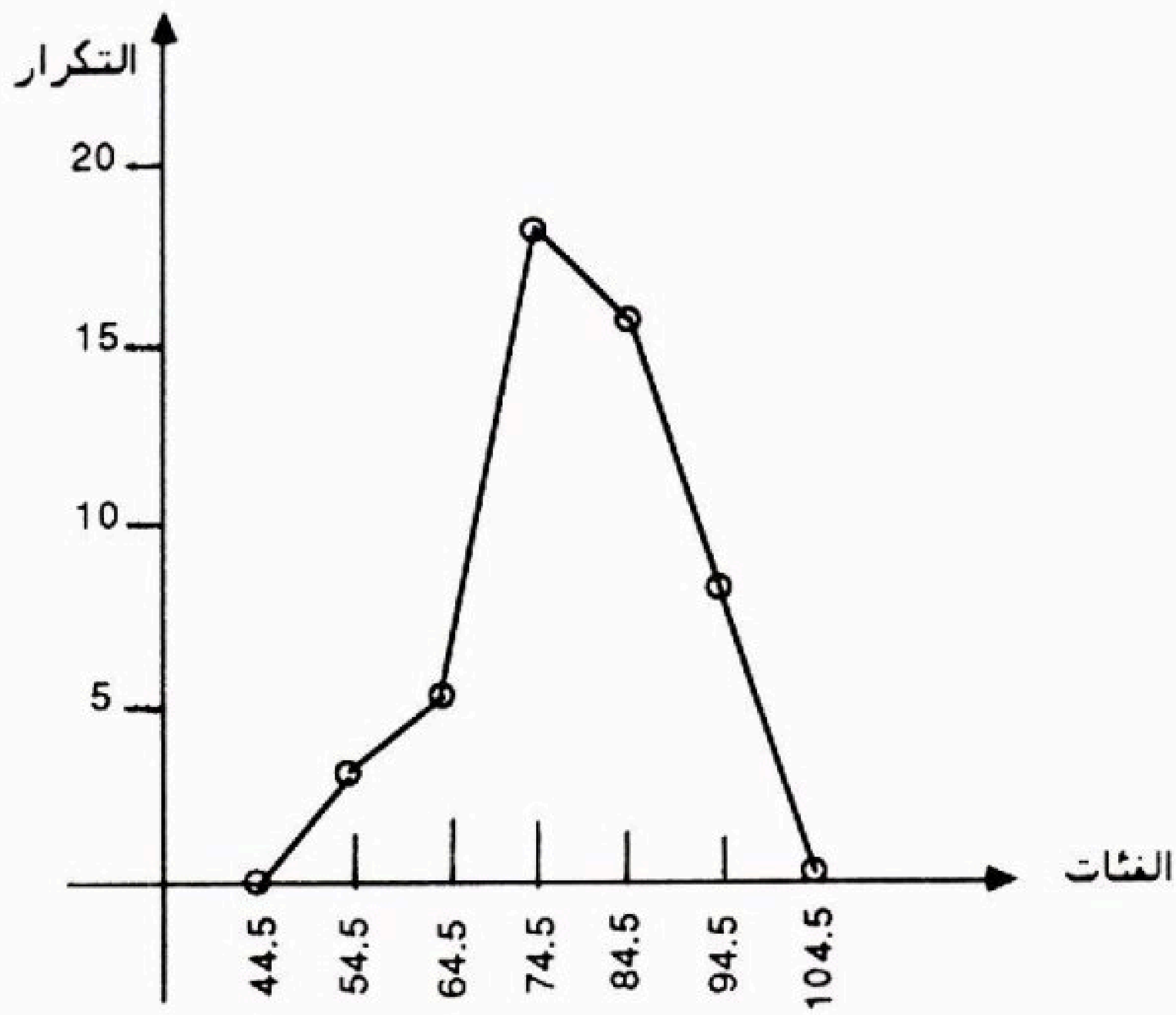
يرسم المضلع التكراري على محورين الأفقي يمثل الفئات والرأسي يمثل التكرار مثل ما ورد شرحه في طريقة رسم المدرج التكراري وبدلاً من رسم مستطيل ارتفاعه يمثل التكرار نضع نقطة واحدة فقط على ارتفاع يمثل التكرار لهذه الفئة وذلك عند منتصف الفئة. ويكرر رسم النقاط لباقي التكرارات. بحيث يكون ارتفاعاتها ممثلة لتكرار تلك الفئات وذلك من منتصفاتها لأننا نفترض انتظام توزيع التكرارات داخل كل فئة. وبعد ذلك نصل بخط مستقيم كل نقطتين متجاورتين. فنحصل على المضلع التكراري. والمضلع التكراري للبيانات في جدول (٢ - ٥) يوضح بالرسم شكل (٢ - ٣) كالتالي:



شكل (٢ - ٣) المضلع التكراري قبل الإغلاق لدرجات الطلاب

ولغلق المضلع التكراري في شكل (٢ - ٣) مع محور الفئات نضع نقطة على محور الفئات يسار الحد الأدنى الحقيقي للفئة الأولى على بعد يساوي نصف طول الفئة ثم نصل بخط مستقيم هذه النقطة بالنقطة التي سبق وضعها في مركز الفئة الأولى. ثم نضع

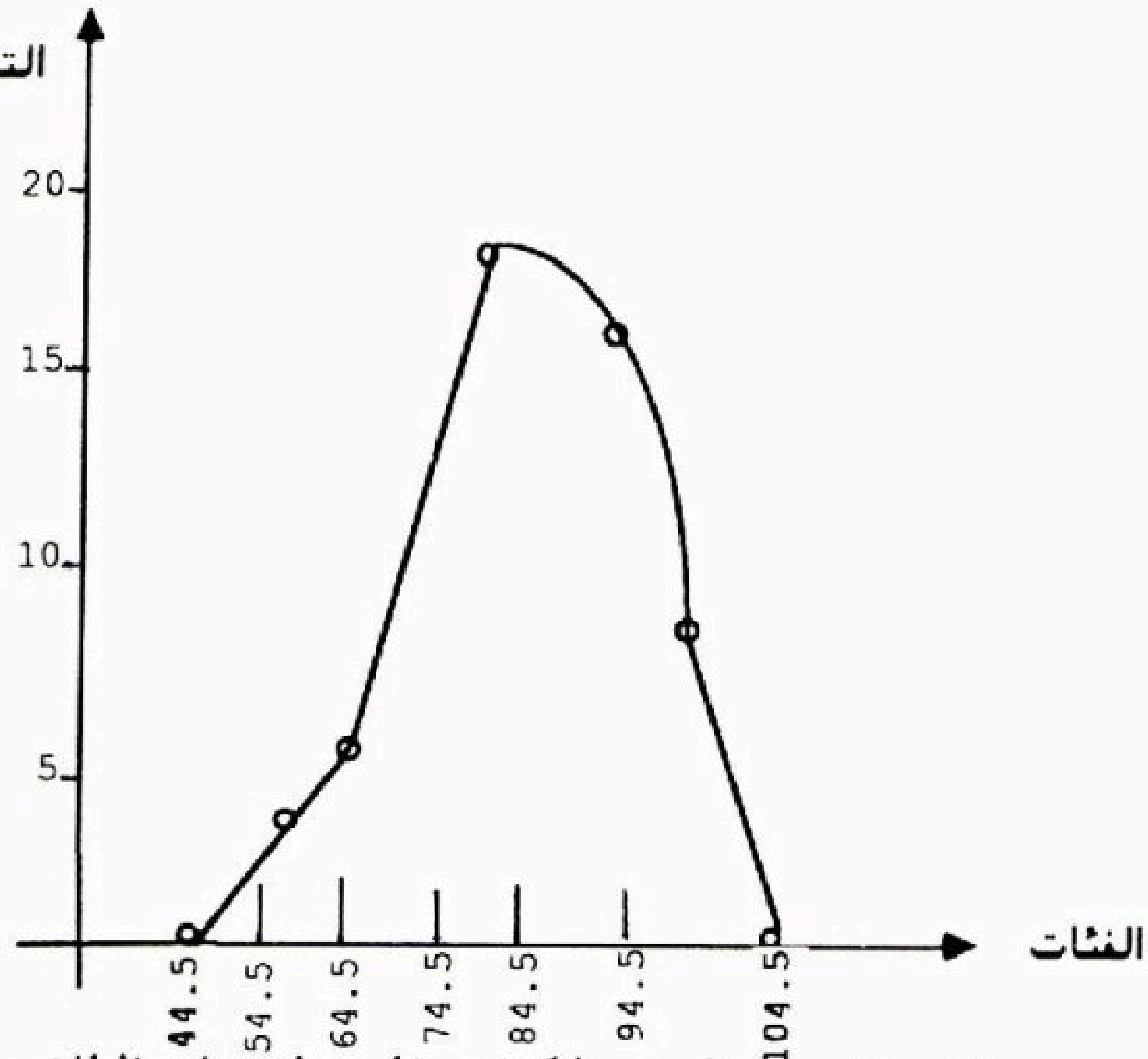
نقـطة على محور الفئـات يمين الحد الأعلى الحقيقي للفئة الأخيرة تبعد مسافة قدرها نصف طول الفئة عن يمين الحد الأعلى الحقيقي للفئة الأخيرة ثم نوصلها بخط مستقيم بالنقطة التي سبق وضعها في منتصف الفئة الأخيرة ولكي يكون المضلع صحيحًا يجب أن يكون مغلقًا ويبين المضلع التكراري المغلق لجدول (٢ ـ ٥) السابق بالشكل (٢ ـ ٤) التالي :



شكل (٢ ـ ٤) المضلع التكراري المغلق لدرجات الطلاب

(٢ ـ ٢ ـ ٣) المنحنى التكراري الممهد (frequency curve)

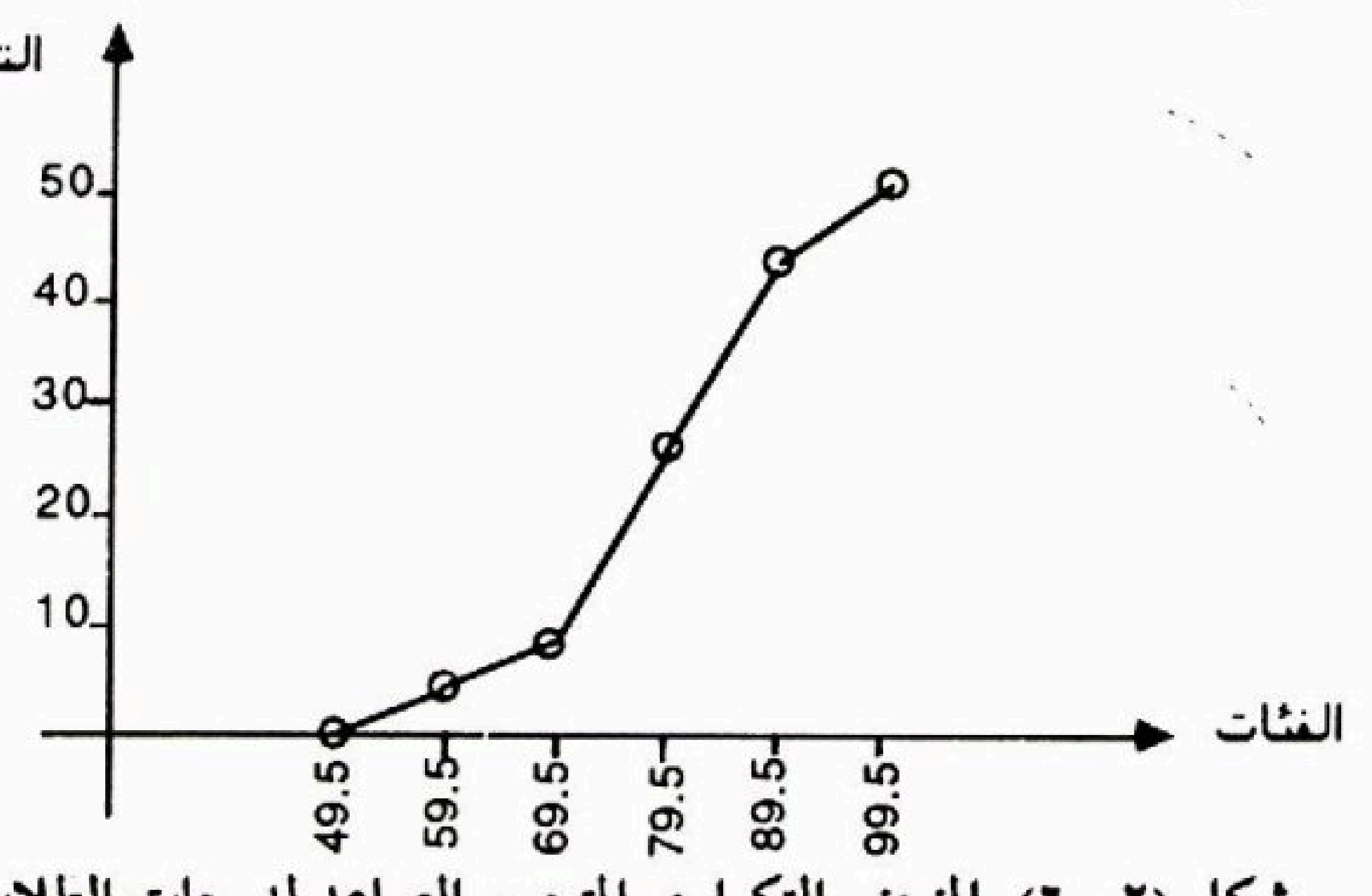
يرسم المنحنى التكراري على محورين متعامدين الأفقي يمثل الفئات والرأسي يمثـل التكرار. ويتم رسم النقاط مثل ما اتبع في المضلع التكراري ويمهد المنحنى التكراري باليد كي يأخذ شكل منحنى إنسيابي حتى لو لزم الأمر عدم المرور ببعض النقاط. المنحنى التكراري الممهد للبيانات في جدول (٢ ـ ٥) كما في شكل (٢ ـ ٥) التالي :



شكل (٢ ـ ٥) المنحنى التكراري الممهد لدرجات الطلاب

(٢ ـ ٢ ـ ٤) المنحنى التكراري المتجمع الصاعد (Ascending cumulative frequency curve)

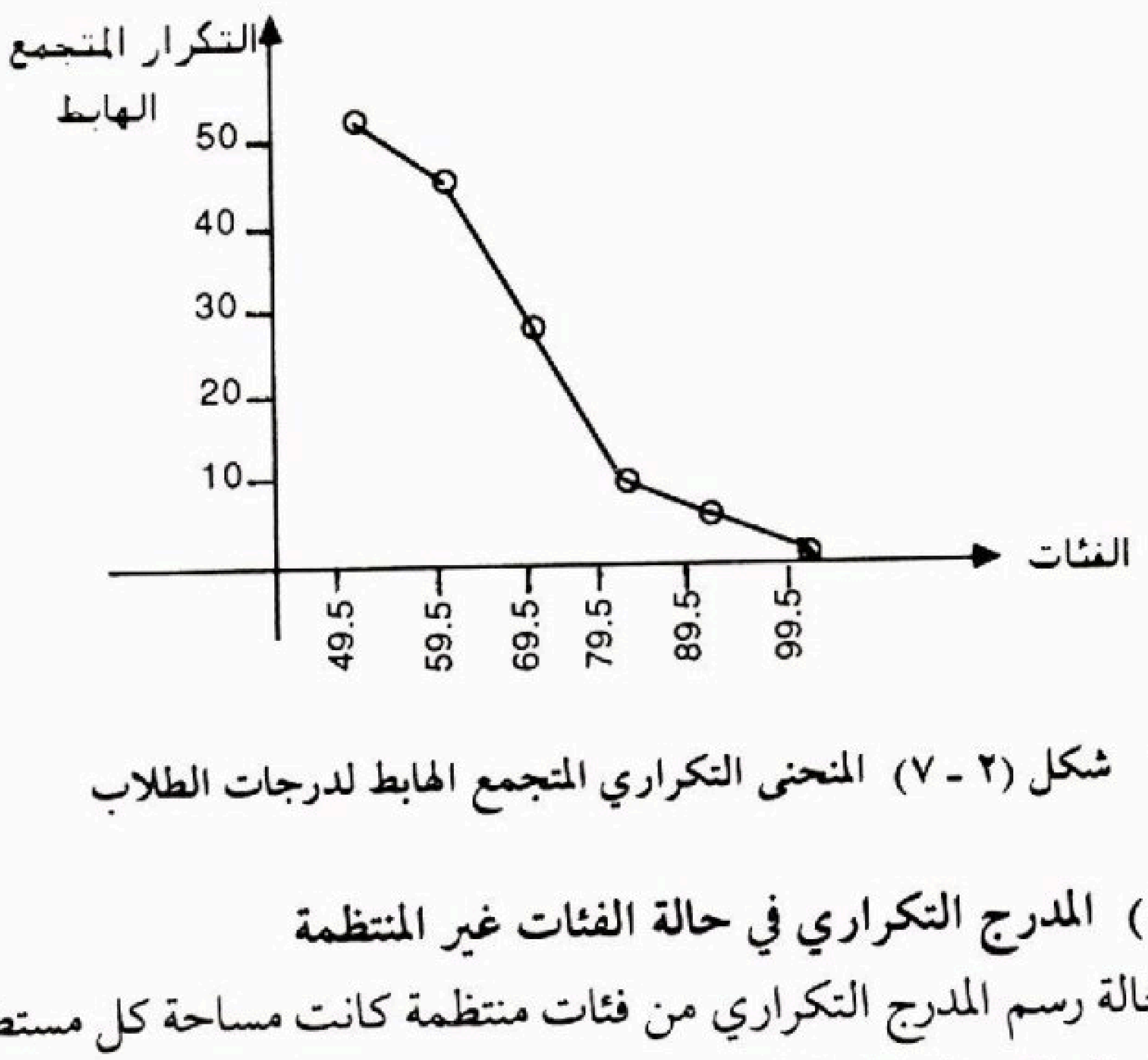
يرسم المنحنى التكراري المتجمع الصاعد على محورين متعامدين الأفقي يمثل الفئـات والرأسي يمثل التكرارات المتجمعة الصاعدة وتوضع النقاط في الرسم أعلى الحدود الدنيا الحقيقية للفئات بحيث يكون الارتفاع ممثلاً للتكرار المتجمع الصاعد. والمنحنى المتجمع الصاعد الممثل في الجدول (٢ ـ ١٠) يوضح بالرسم شكل (٢ ـ ٦) التالي:



شكل (٢ ـ ٦) المنحنى التكراري المتجمع الصاعد لدرجات الطلاب

(٢ ـ ٢ ـ ٥) المنحنى التكراري المتجمع الهابط (descending cumulative frequency curve)

يمثل المنحنى التكراري المتجمع الهابط على محورين متعامدين مثل ما تم بالنسبة للمنحنى المتجمع الصاعد بحيث يكون المحور الأفقي يمثل الحدود الدنيا الحقيقية للفئات والرأسي يمثل التكرارات المتجمعة الهابطة ويمهد المنحنى باليد لنحصل على المنحنى التكراري المتجمع الهابط كما هو موضح بالرسم شكل (٢ ـ ٧) للبيانات في مثال (٢ ـ ٢) وفي جدول (٢ ـ ١١).



شكل (٢ ـ ٧) المنحنى التكراري المتجمع الهابط لدرجات الطلاب

(٢ ـ ٢ ـ ٦) المدرج التكراري في حالة الفئات غير المنتظمة

في حالة رسم المدرج التكراري من فئات منتظمة كانت مساحة كل مستطيل تعبر عن التكرار الواقع في كل فئة. وحيث إن الفئات متساوية في أطوالها فإن المدرج التكراري عبارة عن مستطيلات متلاصقة ومتساوية القاعدة وارتفاعاتها تتناسب مع التكرار. أما إذا كانت الفئات غير متساوية الطول تكون مساحة هذه المستطيلات المتلاصقة غير متناسبة مع التكرار وكذلك ارتفاعاتها لذلك يجب تعديل التكرار قبل رسم المدرج التكراري للفئات غير المتساوية حتى يصبح التكرار المعدل يتناسب مع ارتفاع المستطيل الخاص بالفئة غير منتظمة الطول. وهناك طريقتان لتعديل الجدول

التكراري ، **الطريقة الأولى** : هو أن نقسم التكرار الأصلي لكل فئة على طولها فنحصل على تكرار معدل لجميع الفئات **والطريقة الثانية** : بأن يعدّل تكرار الفئات غير المنتظمة فقط ويترك التكرار للفئات المنتظمة الباقية كما هو ويعدّل التكرار للفئة غير المنتظمة بالعلاقة التالية :

$$\text{التكرار المعدّل} = \frac{\text{التكرار الفعلي للفئة غير المنتظمة} \times \text{طول الفئة المنتظمة}}{\text{طول الفئة غير المنتظمة}}$$

ونوضح ذلك بالمثال التالي .

مثال (٢ ـ ٥)

جدول (٢ ـ ١٦) التالي يبين التوزيع التكراري لفئات غير منتظمة لدرجات الطلاب في مثال (٢ ـ ٢) السابق .

ارسم المدرج التكراري وذلك بعد تعديل التكرارات .

جدول (٢ ـ ١٦) للفئات غير المنتظمة لدرجة الطلاب

التكرار	جدول فئات غير منتظمة
8	50 – 69
18	70 – 79
16	80 – 89
9	90 – 99
50	المجمـوع

الحــــل

نعدّل تكرار الفئة غير المنتظمة (69 – 50) بالعلاقة التالية :

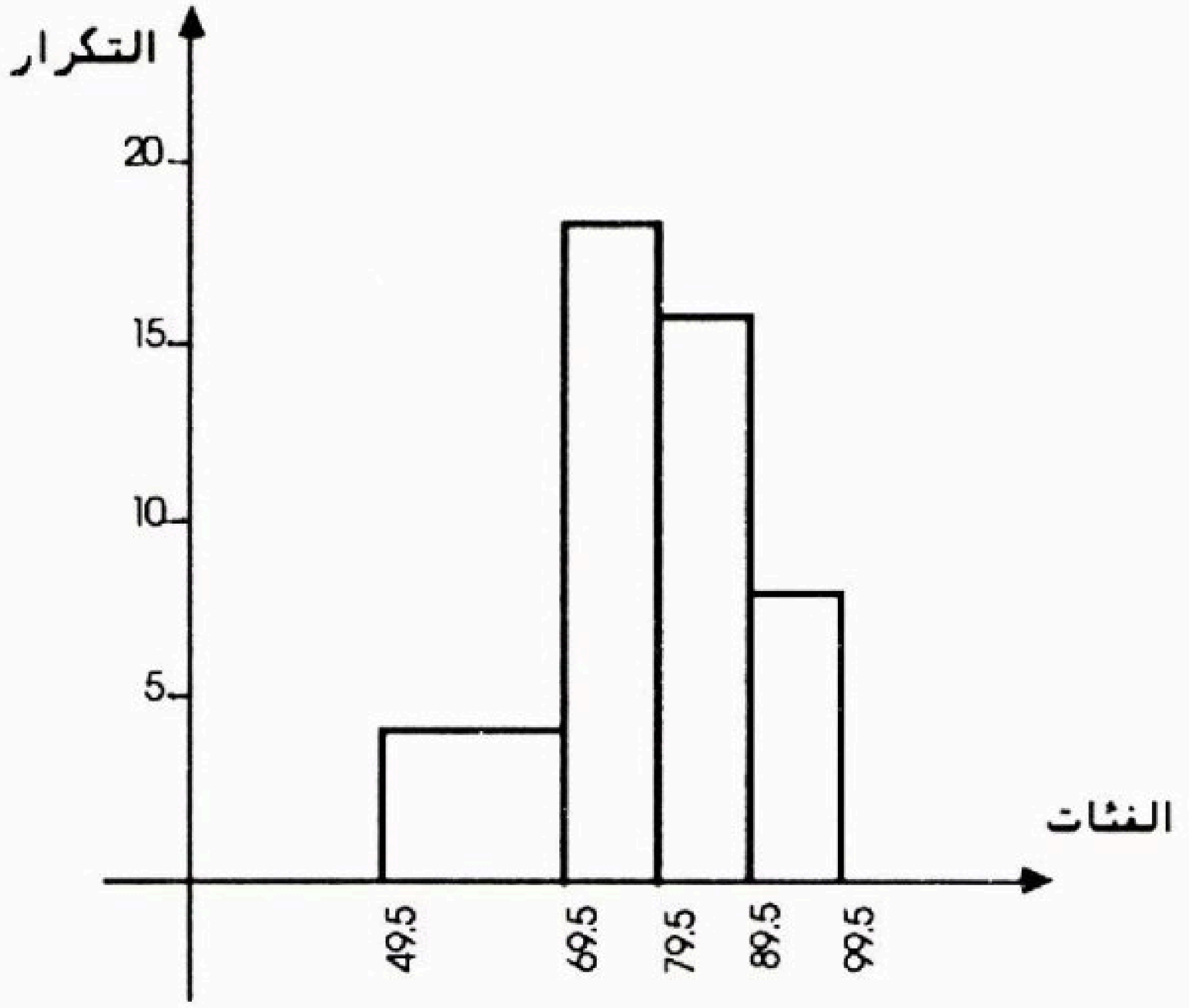
$$\text{التكرار المعدّل للفئة } (50-69) = \frac{\text{تكرار هذه الفئة} \times \text{طول الفئة المنتظمة}}{\text{طول هذه الفئة}} = \frac{10 \times 8}{20} = 4$$

ونوضح ذلك بجدول (٢ ـ ١٧) التالي المشتمل على التكرار المعدل كمايلي :

جدول رقم (٢ ـ ١٧) الفئات المعدّلة لدرجات الطلاب

التكرار المعدل	طول الفئة	التكرار قبل التعديل	حدود الفئات
4	20	8	50 – 69
18	10	18	70 – 79
16	10	16	80 – 89
8	10	6	90 – 99
46		50	المجموع

ويفسر نقص مجموع التكرار المعدل لكبر طول الفئة غير المنتظمة عن طول الفئات المنتظمة ثم نرسم من جدول (٢ ـ ١٧) المدرج التكراري كما في شكل (٢ ـ ٨) التالي :



شكل (٢ ـ ٨) مدرج تكراري يشتمل على فئة غير منتظمة الطول بعد تعديل تكرارها

ملاحظـــة

بما أن الإحـداثي الصادي (المحور الرأسي) يتناسب مع مساحة المستطيل أو التكرار (في حالة المنحنى المتجمع الصاعد) فإن استعمال التكرار الفعلي (الحقيقي) أو النسبي أو المئوي شيء واحد.

(٢ ـ ٢ ـ ٧) بعض الأشكال للمنحنيات التكرارية

قبل التعرض للأشكال البيانية لبعض أنواع المنحنيات التكرارية نعرض بإيجاز بعض خواص المنحنيات بوجه عام:

المنحنى المتماثل

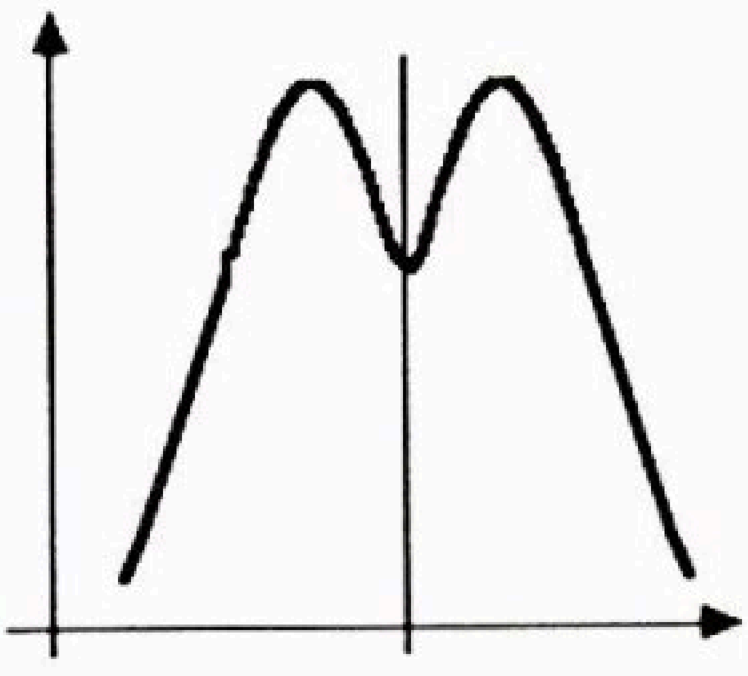
هو المنحنى الذي يتماثل (يتناظر) حول محور يقسمه إلى قسمين متكافئين تمامًا.

المنحنى غير المتماثل

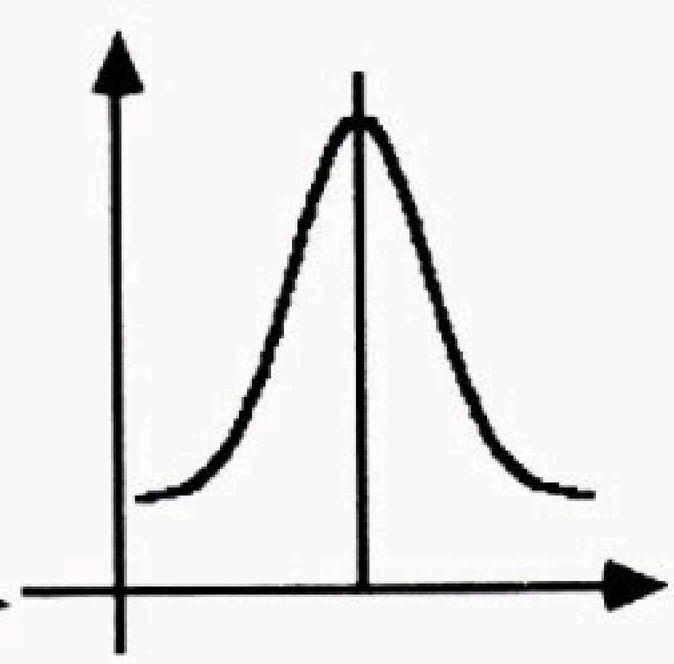
هو المنحنى الذي لا يكون له محور متماثل (تناظر) ويقال له ملتوٍ نحو اليمين أي موجب الالتواء إذا امتد أكثر نحو اليمين وملتوٍ نحو اليسار إذا امتد أكثر نحو اليسار أي سالب الالتواء.

ويوجـد في الحياة العملية كثير من المنحنيات غير المتماثلة وقليل من المنحنيات المتماثلة وسنعرض بعضًا منها كمايلي :

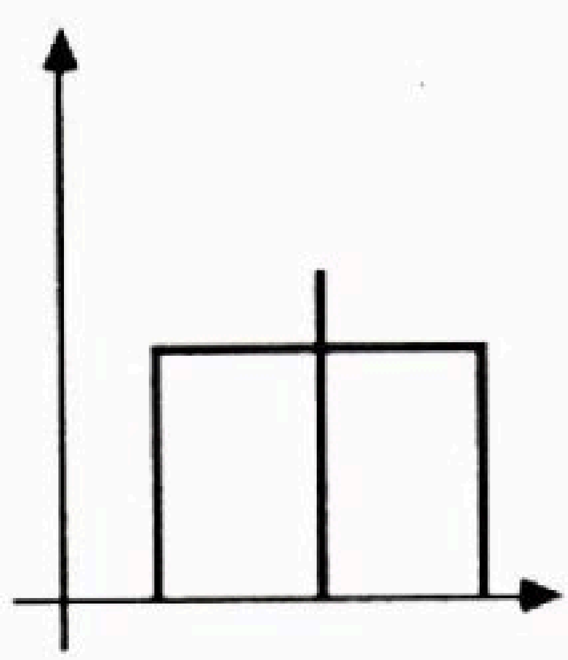
بعض المنحنيات المتماثلة



شكل (٢ - ٩)



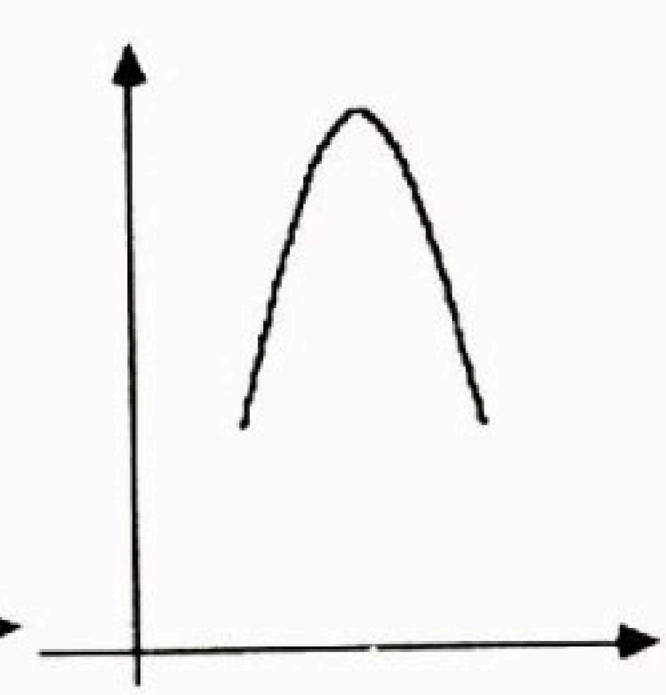
شكل (٢ - ١٠)



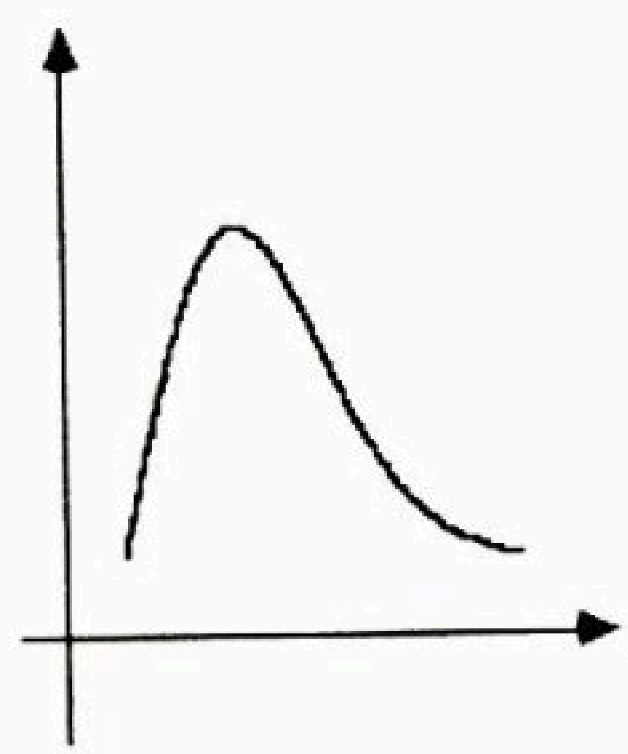
شكل (٢ - ١١)

نلاحظ أن الأشكال الموضحة في (٢ - ٩)، (٢ - ١٠)، (٢ - ١١) تمثل منحنيات متماثلة حول محور.

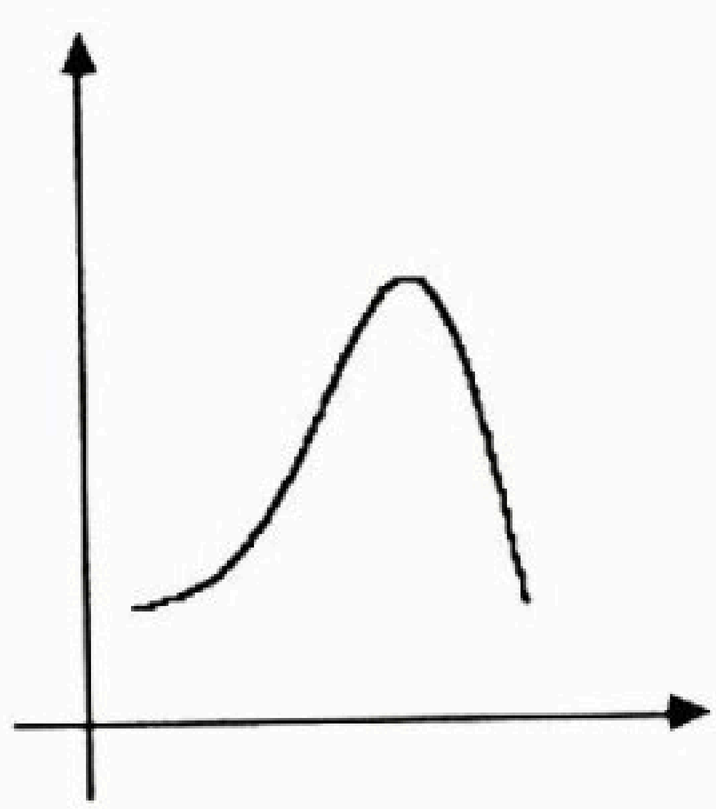
بعض المنحنيات غير المتماثلة



شكل (٢ - ١٢)
معتدل الالتواء



شكل (٢ - ١٣)
ملتوٍ نحو اليمين موجب الالتواء



شكل (٢ - ١٤)
ملتوٍ نحو اليسار سالب الالتواء

(٢ ـ ٣) الرســوم البيانيــة

هنـاك بعض الجـداول الإحصـائية يلزم عرضهـا في شكل رسومات هندسية لتبسيطها وجعل الرؤية للعلاقة بين المتغيرات أكثر سهولة من الجدول من حيث الزيادة أو النقصان لبعض الظواهر الاجتماعية والتعليمية والتجارية والصناعية وغيرها خلال فترة زمنية محددة، ومن أهم هذه الـطرق التي سوف نستعـرضها هي الخط البياني، الأعمدة البيانية، الرسوم الدائرية. وسوف نتناول كل طريقة بالشرح والتفصيل كما يلي:

(٢ ـ ٣ ـ ١) الخط البياني (line graph, or line chart or line diagram)

هو عبارة عن خط منكسر يمثل مسار البيانات الموجودة في الجدول وعادة يستخدم في حالـة البيانات المأخوذة على فترات زمنية. والمحور الأفقي يمثل الزمن (بالسنوات أو الشهور أو الأيام. . . .) والمحور الرأسي يمثل قيم هذه البيانات. والأمثلة على ذلك كثيرة منها على سبيل المثال لا الحصر، تطور التعليم في المملكة العربية السعودية خلال خمس سنوات، أو الاستيراد والتصدير خلال فترة زمنية محددة. . . . الخ.

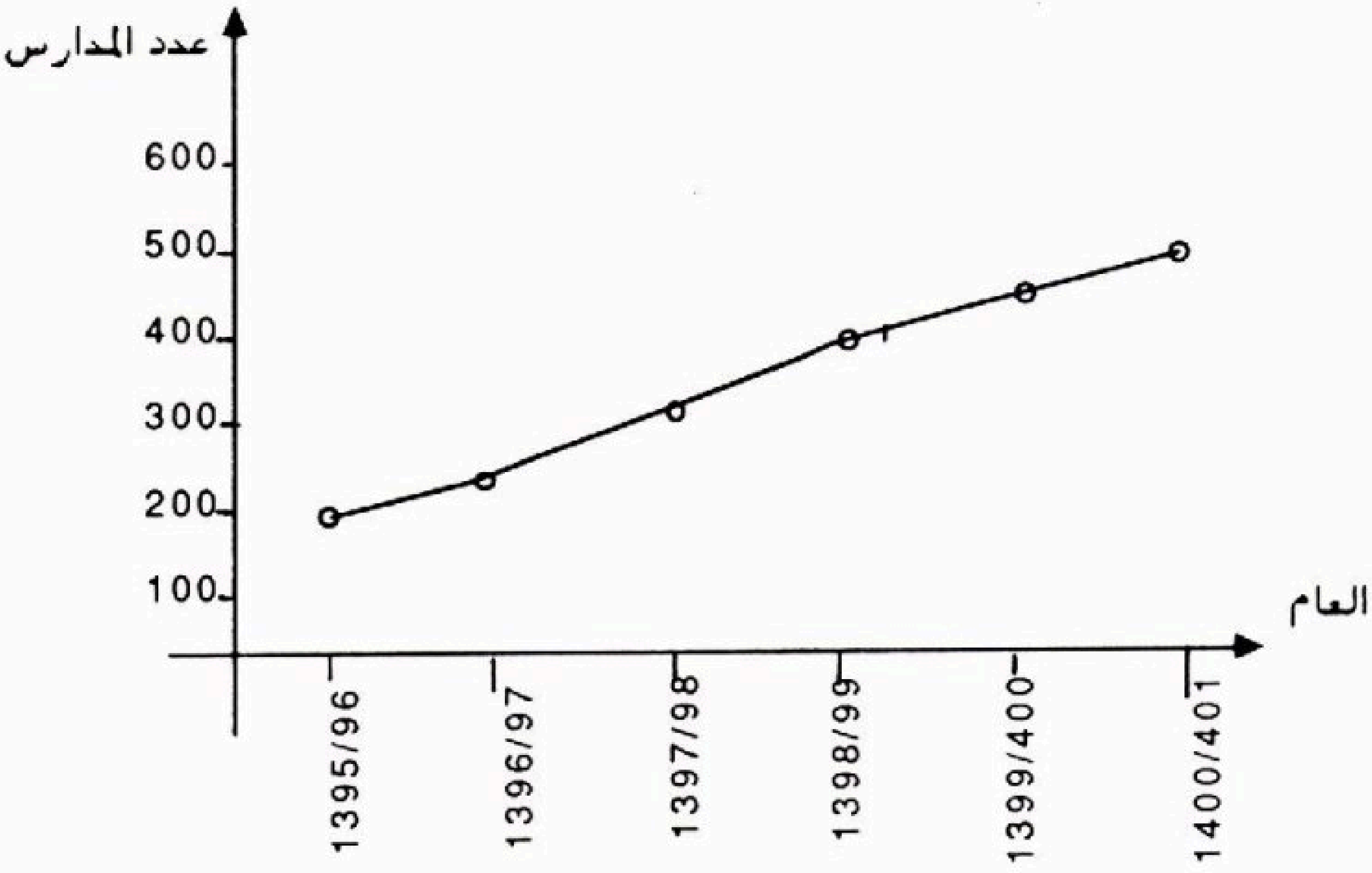
مثال (٢ ـ ٦)

فيما يلي جدول (٢ ـ ١٨) يمثـل عدد المـدارس الثـانـوية في المملكة العربية السعودية بداية من عام 1395/1396 إلى عام 1400/1401 (المصدر: مركز المعلومات الإحصائية)

جدول (٢ ـ ١٨)

السـنـة	1395/96	1396/97	1397/98	1398/99	1399/400	1400/401
عدد المدارس	212	257	331	407	46	513

ويمثل الجدول السابق بالخط البياني كما في شكل (٢ ـ ٥) التالي:



شكل (٢ ـ ١٥): يمثل الخط البياني لعدد المدارس الثانوية بالمملكة وتطورها في فترة من عام ١٣٩٥/ ١٣٩٦هـ إلى عام ١٤٠٠/ ١٤٠١هـ

وإذا كان لدينا أكثر من ظاهرة وقراءات كل منها معطاة في الفترة الزمنية نفسها ويراد المقارنة بينها فإننا نرسم أكثر من خط للظواهر في رسم واحد وكل خط يلوّن بلون خاص أو بخطوط متصلة أو متقطعة ونوضح ذلك بمثال (٢ ـ ٧).

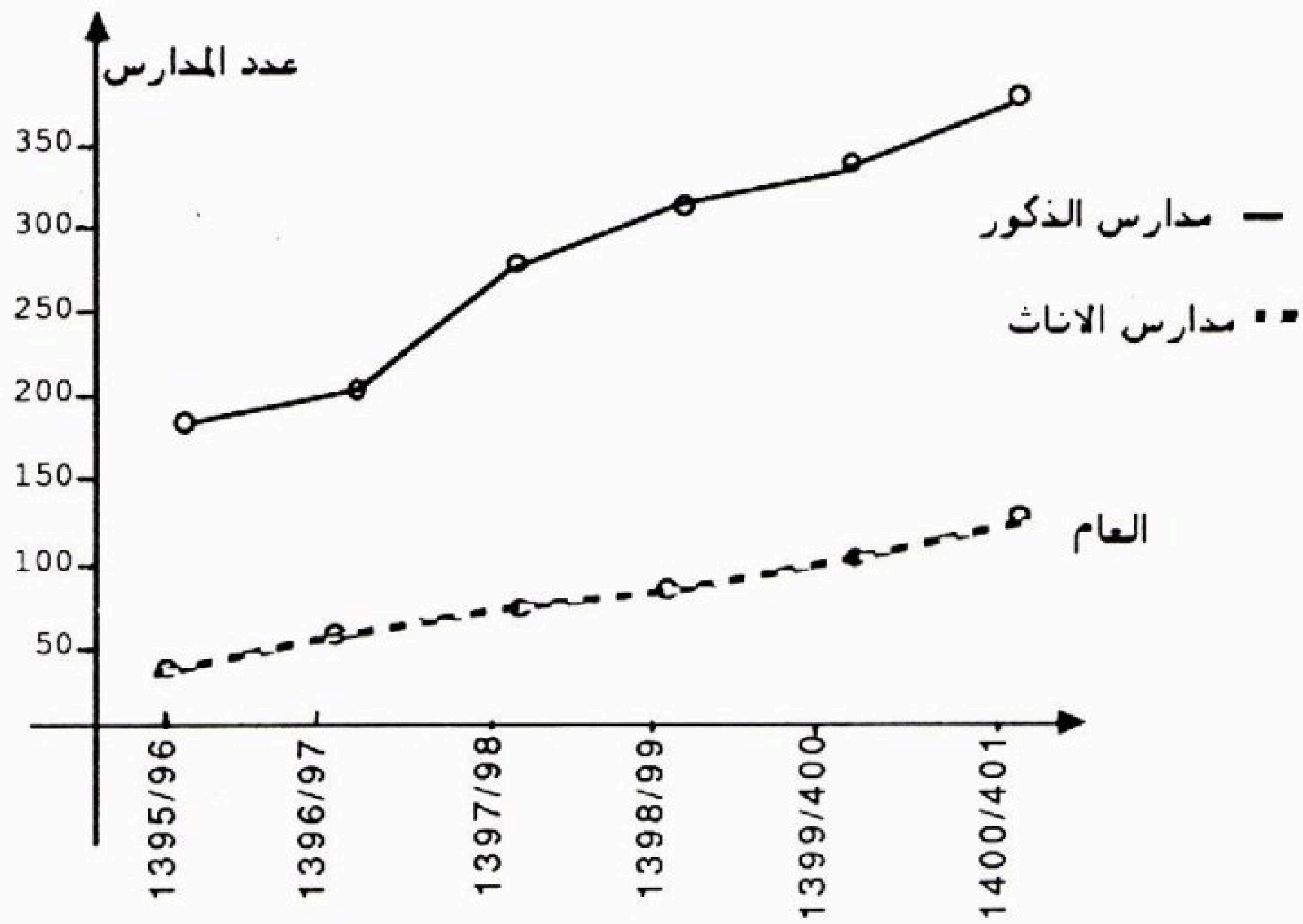
مثال (٢ ـ ٧)

فيما يلي جدول (٢ ـ ١٩) يمثل عدد المدارس الثانوية للذكور والإناث في المملكة العربية السعودية في الفترة من 1395/96 إلى 1400/1401

جدول (٢ ـ ١٩)

السنة	1395/96	1396/97	1397/98	1398/99	1399/400	1400/401
عدد مدارس الذكور	177	209	273	322	343	375
عدد مدارس الاناث	35	48	58	85	113	138

ويمثل الجدول السابق بالخطين البيانيين الآتيين : كما في شكل (٢ ـ ١٦) التالي :



شكل (٢ ـ ١٦) . يمثل الخطين البيانيين لعدد مدارس الذكور وعدد مدارس الإناث .

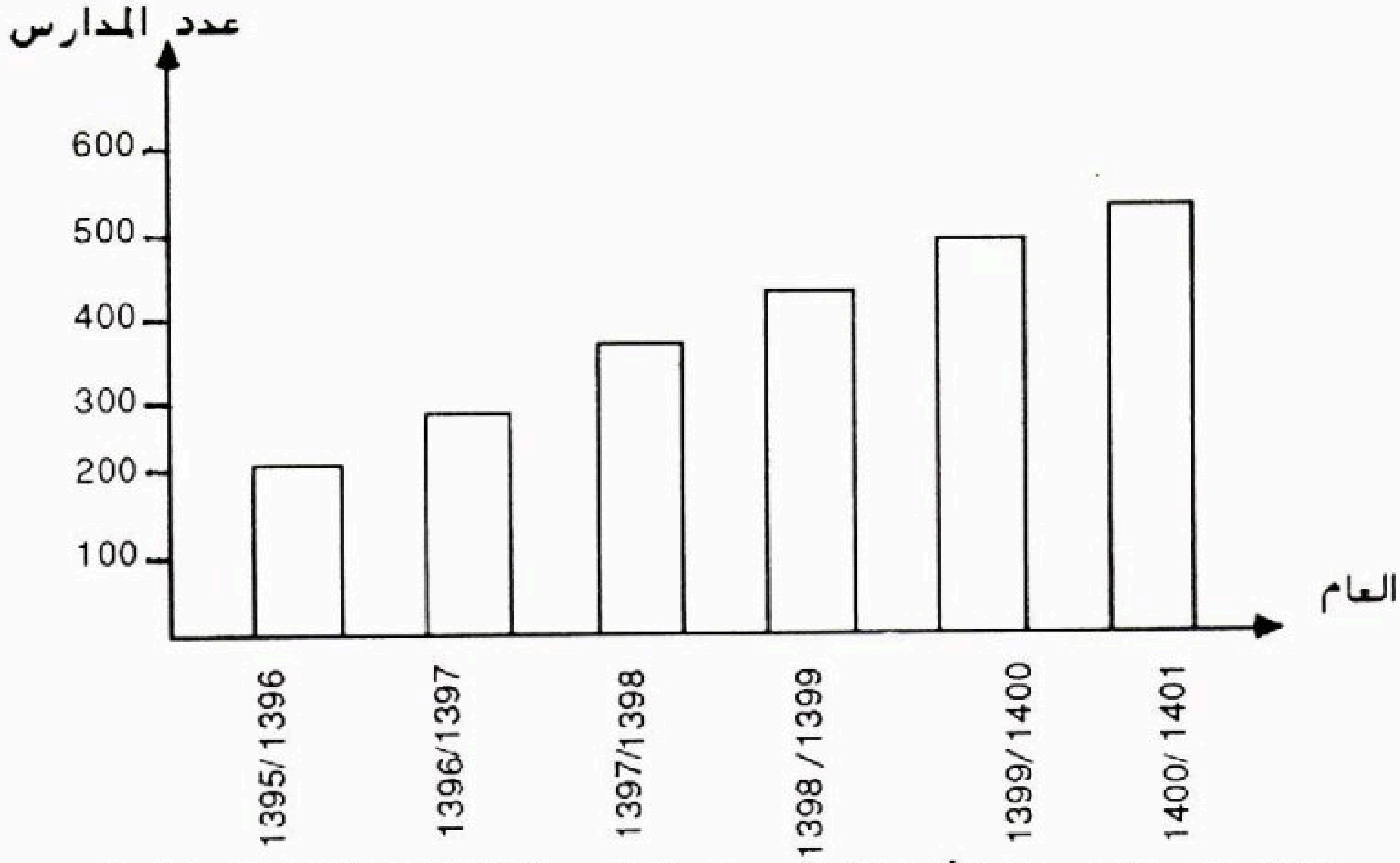
(٢ ـ ٣ ـ ٢) الأعمدة البيانية (**bar graph,** or **bar chart** or **bar diagram**)

وهي عبارة عن مستطيلات رأسية قاعدتها ذات سمك مناسب متساوي وارتفاعاتها تمثل القراءات للظاهرة محل الدراسة وسوف نتكلم بالشرح والتفصيل عن الأعمدة البسيطة والأعمدة المزدوجة والأعمدة المجزأة .

(لاحظ أنها تختلف عن المدرجات التكرارية إذ أنها لا تتقيد بطول الفئة ومساحات المستطيلات ليس لها علاقة بتكرار الفئة)

الأعمدة البيانية البسيطة

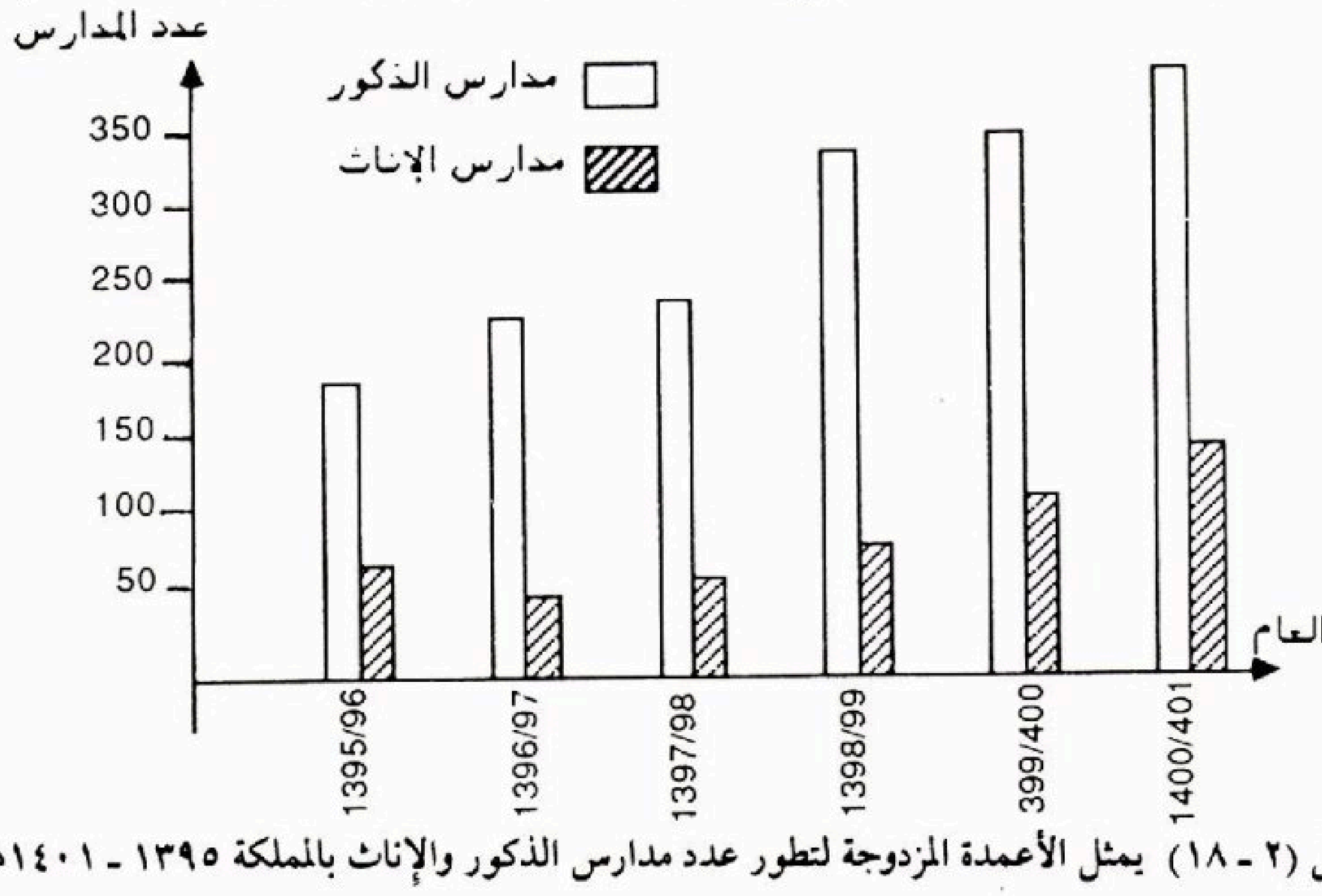
وتستخدم لتمثيل قراءات ظاهرة واحدة وليس من الضروري أن تكون قراءات مقيسة بالنسبة للزمن والأعمدة البيانية للبيانات في مثال (٢ ـ ٦) شكل (٢ ـ ١٧) كالتالي :



شكل (٢ ـ ١٧) يمثل الأعمدة لتطور عدد المدارس في الفترة (١٣٩٥ ـ ١٤٠١هـ)

الأعمدة البيانية المزدوجة

تستخدم الأعمدة المزدوجة عادة إذا أردنا المقارنة بين ظاهرتين ومقارنة التطور بينهما وهي عبارة عن عمودين متلاصقين لكل قراءتين متناظرتين وكل الأعمدة الخاصة بالظاهرة الأولى تلوّن بلون خاص أو تظليل يختلف عن لون وتظليل الظاهرة الثانية حتى يمكن بسهولة المقارنة بينهما ونوضح ذلك في مثال (٢ ـ ٧) كما في شكل (٢ ـ ١٨) التالي :



شكل (٢ ـ ١٨) يمثل الأعمدة المزدوجة لتطور عدد مدارس الذكور والإناث بالمملكة ١٣٩٥ ـ ١٤٠١هـ

الأعمدة البيانية المجزأة

هي عبارة عن أعمدة بيانية بسيطة إلا أن ارتفاعاتها تمثل مجموع القراءات المتناظرة للظاهرتين أو أكثر. ثم يقسّم كل عمود بنسب قراءات الظاهرة وكل ظاهرة تلوّن أو تظلّل بشكل خاص كما يتضح من المثال التالي.

مثال (٢ ـ ٨)

مثل الأعمدة المجزأة للبيانات الواردة في مثال (٢ ـ ٧) السابق.

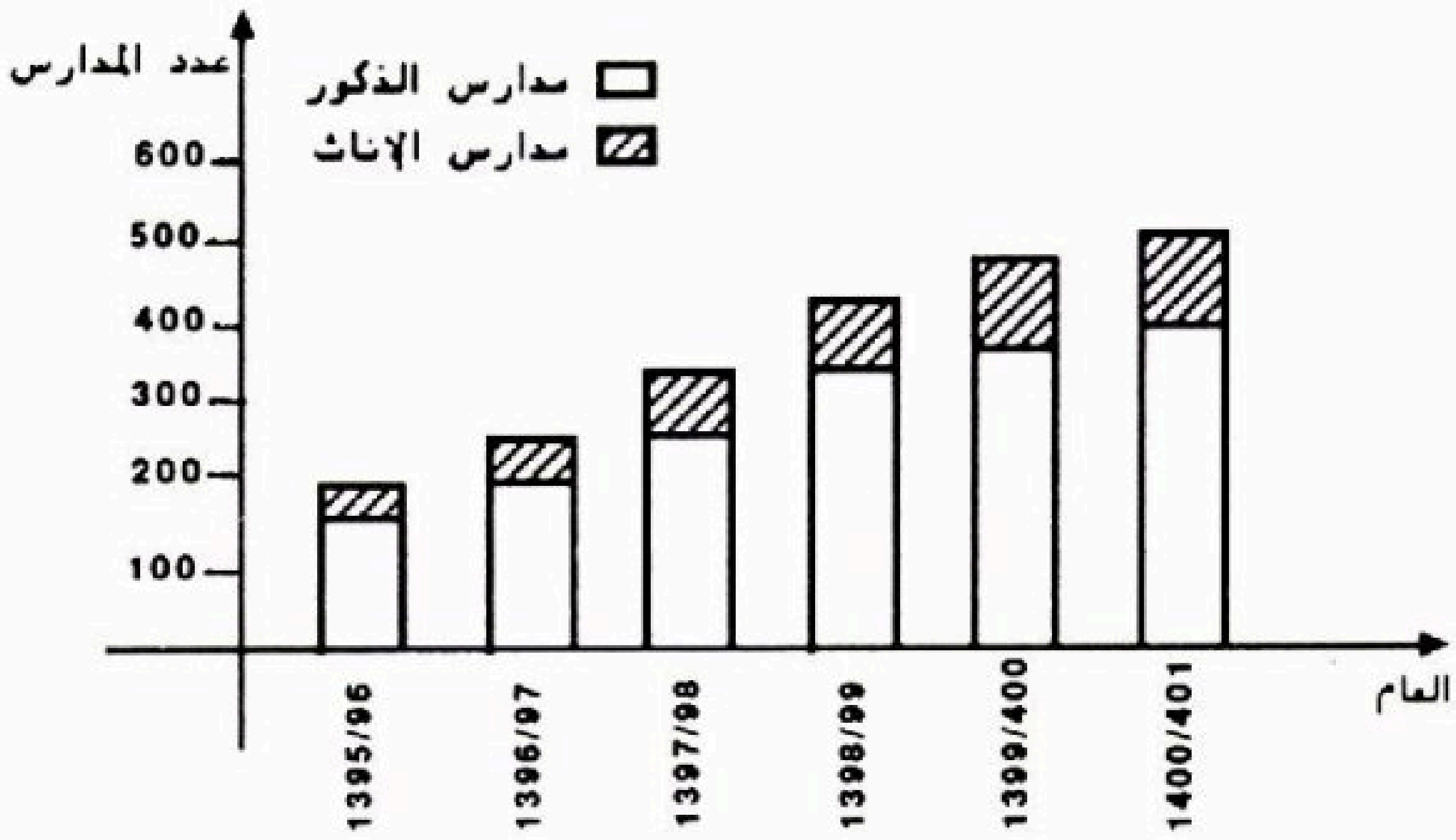
الحــل

ولرسم الأعمدة المجزأة نكون جدول (٢ ـ ٢٠) التالي:

جدول (٢ ـ ٢٠)

السنـة	1395/96	1396/97	1397/98	1398/99	1399/400	1400/401
عدد مدارس الذكور	177	209	273	322	343	375
عدد مدارس الإناث	35	48	58	85	113	138
المجمـوع	212	357	331	407	456	513

ومن الجدول (٢ ـ ٢٠) السابق يمكن رسم الأعمدة المجزأة كما هو موضح في شكل (٢ ـ ١٩) التالي:



شكل (٢ ـ ١٩) يمثل الأعمدة المجزأة لتطور عدد مدارس الذكور والإناث بالمملكة (١٣٩٥ـ١٤٠١هـ)

(٢ ـ ٣ ـ ٣) الرسوم الدائرية (pie chart)

وهي عبارة عن دائرة تقسم إلى قطاعات زواياها المركزية تتناسب مع القراءات ويمكن حساب الزاوية الخاصة بقطاع يمثل قراءة من القراءات من القانون الآتي:

$$\text{الزاوية المركزية لقطاع ممثل لقراءة معينة} = \frac{360^\circ}{\text{مجموع القراءات}} \times \text{القراءة نفسها}$$

مثال (٢ ـ ٩)

فيما يلي جدول (٢ ـ ٢١) يمثل مساحات القارات للعالم، مثلها بالرسوم الدائرية.

جدول (٢ ـ ٢١)

القـــارة	المساحة بالمليون كم٢
أفريقيا	30.3
آسيـا	47.4
أوربـا	4.9
أمريكا الشمالية	24.3
أستراليا ونيوزلندا	8.5
أمريكا الجنوبية	17.9

الحـــل

نكوّن جدول الحل (٢ ـ ٢٢) كالآتي :

جدول (٢ ـ ٢٢)

القــارة	المساحـة	الزاوية المركزية
أفريقيـا	30.3	81.83 = 82
آسيـــا	47.4	128.01 = 128
أوربـــا	4.9	13.23 = 13
أمريكا الشمالية	24.3	65.63 = 66
أستراليا ونيوزلندا	8.5	22.96 = 23
أمريكا الجنوبية	17.9	48.34 = 48
المجمـــوع	133.3	360

من الجدول (٢ ـ ٢٢) السابق يمكن رسم القطاعات الدائرية كما هو موضح في الشكل (٢ ـ ٢٠) التالي .

شكل (٢ ـ ٢٠) يمثل مساحات قارات العالم بالرسوم الدائرية

وأحيانًا يكون لدينا النسب المئوية للقراءات أو حسابها كما يلي :

$$\text{النسبة المئوية للقراءة} = \frac{\text{القـــــراءة}}{\text{مجموع القراءات}} \times 100$$

فإن هذه النسب المئوية يمكن تمثيلها بالرسوم الدائرية وتحسب الزاوية المركزية للقطاع الـذي يمثـل أي نسبـة مئـوية، بضرب هذه النسبـة في (3.6) كما هو موضح بمثال (٢ ـ ١٠) التالي.

مثال (٢ ـ ١٠)

فيما يلي جدول (٢ ـ ٢٣) يمثل النسب المئوية لميزانية جهات التعليم المختلفة بالمملكة العربية السعودية (١٤٠٠ / ١٤٠١هـ).

جدول (٢ ـ ٢٣)

الجهـــــة	النسبـــة المئويـــة
وزارة المعـــارف	45.3
وزارة التعليم العالي والجامعات	35.3
الرئاسـة العامـة لتعليـم البنـــات	19.4

مثِّل هذه البيانات السابقة بواسطة الرسوم الدائرية.

الحـــل

فيما يلي (جدول ٢ ـ ٢٤) الذي يشتمل على الزوايا المركزية للقطاعات الدائرية.

جدول (٢ ـ ٢٤)

الجهـــة	النسبة المئوية	الزاويـة المركزيـة
وزارة المعارف	45.3	3.6X45.3 = 163.08 = 163
وزارة التعليم العالي والجامعات	35.3	3.6X35.3 = 127.08 = 127
الرئاسة العامة لتعليم البنات	19.4	3.6X19.4 = 69.84 = 70
المجمـــوع	100	360

من الجدول (٢ ـ ٢٤) السابق يمكن رسم القطاعات الدائرية للنسبة المئوية لميزانية جهات التعليم المختلفة بالمملكة كما هو موضح في الشكل (٢ ـ ٢١) التالي:

شكل (٢ ـ ٢١) يمثل النسب المئوية لميزانية جهات التعليم المختلفة بالمملكة للعام (١٤٠٠/ ١٤٠١هـ)

(٢ ـ ٣ ـ ٤) أشكال الجذع والورقة البيانية (Stem - and - leaf diagrams)

تعتبر أشكال الجذع والورقة البيانية إحدى الطرق الإحصائية الاستكشافية التي تساعد على تكوين فكرة واسعة عن المدى الذي تغطيه البيانات (أو المشاهدات)، وكيفية تمركزها وكشف أي فجوات في مجال الظاهرة المقاسة عليها البيانات، وإيضاح

أي قيم متطرفة في البيانات. كما أن أشكال الجذع والورقة تستخدم الأرقام الفعلية للبيانات في عرض تلك الأشكال وبهذا لا نفقد القيم الحقيقية، بعكس الطرق البيانية الأخرى، ولكن هذه الأشكال تشترك مع بقية الطرق البيانية الأخرى في كونها لا تحافظ على الترتيب الأصلي للبيانات (وهذه خاصية سيئة إذا كان الترتيب مهمًا).

في شكل الجذع والورقة البياني كل مشاهدة تمثل على شكل ورقة وجذع أو كما يقال ورقة على جذع. الورقة تتشكل من آخر رقم على اليمين من المشاهدة والجذع يتكون من بقية الأرقام فمثلاً للمشاهدة 154 الورقة هي الرقم 4 والجذع هو الرقم 15 . بالنسبة للأرقام التي تحوي فاصلة عشرية مثل الرقم 7.8 تؤخذ الورقة على أنها الرقم 8 والجذع إما يوضح على الشكل .7 أو 7 فقط.

يشكل الجذع والورقة برسم خط رأسي وتوضع الأرقام المكونة للجذع على يسار الخط وتوضع الأرقام المكونة للورقة على يمين الخط أمام الرقم الذي يمثل جذعها مباشرة فمثلاً للبيانات التالية التي تمثل درجات طلاب أحد شعب ١٠١ إحص في سنة ماضية.

28, 46, 49, 42, 58, 59, 53, 50, 51, 66, 62, 64, 67, 69,

63, 68, 67, 69, 75, 70, 78, 75, 74, 84, 85, 88, 93, 99

نلاحظ أن المشاهدات تتغير بين 28 و 99 وعليه نشكل الجذع من الأرقام 2 إلى 9 شكل (٢ ـ ٢٢)

نستعرض المشاهدات بالترتيب فنجد أن المشاهدة الأولى هي 28 فنضع الرقم 8 أمام الرقم 2 الموجود في الجذع،شكل (٢ ـ ٢٣) المشاهدة التالية هي 46 نضع الرقم 6 أمام الرقم 4 الموجود في الجذع،شكل (٢ ـ ٢٤) المشاهدة التالية لها هي 49 فنضع الرقم 9 أمام الرقم 4 الموجود في الجذع،شكل (٢ ـ ٢٥) وهكذا نستمر حتى تنتهي جميع المشاهدات ويكون الناتج هو شكل الجذع والورقة كما في شكل (٢ ـ ٢٦).

2	
3	
4	
5	
6	
7	
8	
9	

شكل (٢ ـ ٢٢)

2	8
3	
4	
5	
6	
7	
8	
9	

شكل (٢ ـ ٢٣)

2	8
3	
4	6
5	
6	
7	
8	
9	

شكل (٢ ـ ٢٤)

2	8
3	
4	69
5	
6	
7	
8	
9	

شكل (٢ ـ ٢٥)

2	8
3	
4	692
5	89301
6	624793879
7	50854
8	458
9	39

شكل (٢ ـ ٢٦) يمثل الجذع والورقة لدرجات الطلاب في مقرر ١٠١ إحص

مثـال (٢ ـ ١١)

أرسم عرض الجذع والورقة للبيانات التالية :

35.4, 36.8, 32.9, 38.9, 37.5, 37.5, 36.7, 36.4, 37.9, 38.6,

, 32.9, 33.4, 37.2, 32.9, 32.0, 33.2, 34.2, 35.7, 35.7, 35.7,

34.4, 36.1, 35.8, 34.5, 36.7, 38.0, 38.7, 34.6, 32.1, 37.2

الحــل

نلاحظ أن المشاهدات تتغير بين 32.9 و 38.9 وعليه نشكل الجذع من الأرقام 32.0 إلى 38.0 ثم نستعرض الأرقام بعد ذلك فنحصل على عرض الجذع والورقة كما في شكل (٢ ـ ٢٧) التالي :

32	9	9	9	0	1
33	4	2			
34	2	4	5	6	
35	4	7	7	7	8
36	8	7	4	1	7
37	5	5	9	2	2
38	9	6	0	7	

شكل (٢ ـ ٢٧) يمثل الجذع والورقة للبيانات في مثال (٢ ـ ١١)

مثـال (٢ ـ ١٢)

أرسم عرض الجذع والورقة للبيانات التالية :

2.7, 2.0, 3.6, 5.0, 6.5, 6.1, 5.9, 5.0, 6.5, 7.4, 8.2, 3.9,

4.1, 4.5, 5.5, 3.8, 4.8, 5.6, 6.3, 5.9, 8.7, 5.3, 5.7, 5.7

الحـــل

نلاحظ أن المشـاهدات تتغير بين 2.0 و 8.7 ومن ثم نشكل الجذع من الأرقام 2.0 إلى 8.0 ثم نستعـرض الأرقـام بعـد ذلك فنحصل على عرض الجذع والورقة كما في شكل (٢ ـ ٢٨) التالي :

2	7 0
3	6 9 8
4	1 5 8
5	0 9 0 5 6 9 3 7 7
6	5 1 5 3
7	4
8	2 7

شكل (٢ ـ ٢٨) يمثل الجذع والورقة للبيانات في مثال (٢ ـ ١٢)

مثـال (٢ ـ ١٣)

أرسم عرض الجذع والورقة للبيانات التالية :

8, 23, 18, 22, 22, 15, 21, 23, 25, 18, 24,

22, 21, 37, 19, 22, 22, 12, 27, 16, 26, 32

الحـــل

نلاحظ أن المشـاهدات تتغير بين 8 و 37 ومن ثم نشكل الجذع من الأرقام 0 فنحصل على عرض الجذع والورقة كما في شكل (٢ ـ ٢٩) التالي :

0	8
1	8 5 8 9 2 6
2	3 2 2 1 3 5 4 2 1 2 2 7 9
3	7 2

شكل (٢ ـ ٢٩) يمثل الجذع والورقة للبيانات في مثال (٢ ـ ١٣)

(٢ ـ ٤) تمـــارين

(١) فيما يلي بيان بأعضاء هيئة التدريس بجامعة الرياض (المصدر: ربع قرن في حياة جامعة الرياض) من عام ١٣٩٥/١٣٩٦هـ وحتى عام ١٤٠٠/١٤٠١هـ.

العام الدراسي / الجنسية	1395/96	1396/97	1397/98	1398/99	1399/400	1400/401
سعـــودي	220	269	345	370	372	491
غير سعودي	593	667	763	902	962	1152

والمطلوب تمثيل هذه البيانات باستخدام:

(أ) ـ الخط البياني.

(ب) ـ الأعمدة البيانية البسيطة ـ المزدوجة ـ المجزأة.

(٢) الجدول الآتي يوضح تطور أعداد خريجي الجامعة من السعوديين حسب الجنس (المصدر: ربع قرن في حياة جامعة الرياض) من عام ١٣٩٥/١٣٩٦هـ حتى عام ١٤٠٠/١٤٠١هـ

العام الدراسي / الجنـس	1395/96	1396/97	1397/98	1398/99	1399/400	1400/401
خـــريـــج	756	722	946	929	984	963
خـــريـجـة	13	46	76	106	205	257

مثّل هذه البيانات باستخدام :

(ا) - الخط البياني .

(ب) - الأعمدة البيانية المختلفة .

(جـ) - الرسوم الدائرية .

(٣) البيانات الآتية تمثل أعداد الطلبة السعوديين الملتحقين بالجامعة حسب الجنس من عام ١٣٩٥/١٣٩٦هـ وحتى عام ١٤٠٠/١٤٠١هـ

العام الدراسي / الجنس	1395/96	1396/97	1397/98	1398/99	1399/400	1400/401
طالب	5703	5795	6880	7925	8477	9295
طالبة	807	1022	1674	2310	2349	2511

مثّل هذه البيانات بطريقتين مناسبتين ومختلفتين .

(٤) اثبت أن المساحة الكلية للمستطيلات في المدرج التكراري تساوي المساحة الكلية المحصورة بين المضلع التكراري المغلق مع المحور الأفقي (السينات) .

(٥) أربع قطع معدنية من الريالات رميت 100 مرة وفي كل مرة سجل عدد الصور فكانت كالآتي :

عدد الصور	0	1	2	3	4
عدد الرميات	11	23	32	25	9

(i) ارسم هذه البيانات بتمثيل بياني مناسب .
(ii) كون جدولاً تظهر فيه النسب المئوية للرميات التي تظهر بها عدد الصور أقل من 0, 1, 2, 3, 4 .
(iii) ارسم بيانات الجدول الذي حصلت عليه في (ii) .

(٦) فيما يلي أوزان 80 فأرًا من فئران التجارب بالجرام وذلك عند دراسة نقص الفيتامين .

132	125	117	124	108	112	110	127	96	129
130	122	118	114	103	119	106	125	114	100
125	128	106	111	116	123	119	114	117	143
136	92	115	118	121	137	139	120	104	125
119	115	101	129	87	108	110	133	135	126
127	103	110	126	118	82	104	137	120	95
146	126	119	119	105	132	126	118	100	113
106	125	117	102	146	129	124	113	95	148

(i) كون جدول التوزيع التكراري مستخدمًا أطوال الفئات الآتية :
80 – 89 , 90 – 99 , 100 – 109 , ... , 140 – 149
(ii) ارسم المدرج التكراري والمضلع التكراري .
(iIi) ارسم المدرج التكراري النسبي والمضلع التكراري النسبي .
(iv) ارسم المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط لهذه البيانات .
(v) أوجد عدد الفئران التي تقل أوزانها عن 125 جرامًا .

(٧) عند دراسة الحالة التعليمية لمجموعة من الزوجات كانت لدينا النتائج التالية :
ابتـدائية ـ أمية ـ ثانـوية ـ ابتـدائية ـ ثانوية ـ متوسطة ـ متوسطة ـ أمية ـ أمية ـ
ابتدائية ـ أمية ـ جامعية ـ جامعية ـ أمية ـ ابتدائية ـ متوسطة ـ متوسطة ـ أمية ـ

ابتـدائية ـ ثانوية ـ أمية ـ ابتدائية ـ جامعية ـ متوسطة ـ ثانوية ـ أمية ـ ثانوية ـ أمية ـ ابتدائية ـ ثانوية ـ جامعية .

(أمية تعني لا تقرأ ولا تكتب)

مثِّل هذه البيانات بطريقة مناسبة .

(٨) فيما يلي درجات عدد من الطلبة :

68	72	64	82	71	66	60	40	98	44
69	79	68	65	60	88	78	77	69	55
62	55	70	83	75	61	66	71	64	62
50	60	66	67	62	74	62	61	72	57

(i) أوجد جدول التوزيع التكراري لهذه الدرجات مستخدمًا الفئات

$40-49$, $50-59$, ... , $90-99$

(ii) ارسم المـدرج التكـراري والمنحنى التكـراري ثم أوجـد مسـاحة المدرج التكراري والمساحة المحصورة بين المضلع التكراري ومحور السينات وقارن بينهما .

(iii) ارسم المنحنى المتجمع الصاعد النسبي والمنحنى الهابط النسبي .

(iv) إذا علم أن :

التقـديـر	الـدرجـــات
هـ	0 - 59
د	60 - 69
جـ	70 - 79
ب	80 - 89
أ	90 - 99

أوجد جدول توزيع التقديرات لدرجات الطلاّب .

(٩) فيما يلي أجور 70 عاملاً في إحدى المؤسسات بالريال في اليوم الواحد .

فئات الأجور	50 – 59	60 – 69	70 – 79	80 – 89	90 – 99	100 – 119	120 – 129
عدد العمال	8	10	16	15	10	8	3

(i) ارسم المضلع التكراري لهذه البيانات .
(ii) ارسم المنحنى التكراري والمنحنى المتجمع الصاعد لهذه البيانات .
(iii) ارسم المنحنى المتجمع الهابط لهذه البيانات .

(١٠) فيما يلي أوزان عدد من الأطفال حديثي الولادة بالرطل

9.0	6.5	9.5	5.1	4.8	8.8	6.5	9.5
7.7	6.9	6.6	6.0	7.9	7.7	6.9	6.6
5.8	7.1	6.8	8.4	6.9	5.8	7.1	6.8
8.6	9.8	3.8	7.4	7.2	8.6	9.8	3.8
10.3	7.4	5.7	4.5	7.7	10.3	7.4	5.7
9.4	7.8	8.7	5.8	8.6	9.4	7.8	8.9
8.8	9.4	6.0	5.9	7.4	8.8	9.4	6.0
7.2	10.5	9.4	7.4	8.9	7.2	10.5	8.4
10.4	7.8	5.0	4.6	8.0	10.2	7.8	5.0

مثِّل هذه البيانات باستخدام طريقة «الجذع والورقة» (stem and leaf) بمدى رطل واحد ثم ظلل شكل «الجذع والورقة» الناتج .

(١١) إذا أعطيت الجدول التكراري الآتي لمراكز فئات الأقطار لستين شجرة من نوع شائع ـ محسوبة بالأقدام ـ فارسم مدرجها التكراري مبينًا حدود الفئة.

القطر بالقدم	1	2	3	4	5	6	7	8	9	10	11	12
عدد الأشجار	1	7	12	16	10	4	5	2	1	1	0	1

(١٢) إذا أعطيت الجدول التكراري الآتي لأطوال ألف طالب مقاسة بالسنتيمترات فارسم مدرجها التكراري مبينًا حدود الفئات الحقيقية.

طول الطالب	155 – 157	158 – 160	161 – 163	164 – 166	167 – 169	170 – 172	173 – 175	176 – 178
عدد الطلاب	4	10	77	235	368	220	80	6

(١٣) فيما يلي بيانات عن وقت السفر إلى العمل ومنه بالساعات في اليوم للعاملين في الطيران. ارسم المدرج التكراري، بفرض اتصال الزمن.

التكرار	الفئات
80	أقل من ساعة
42	من 1 إلى 2
7	من 2 إلى 3
4	من 3 إلى 4
3	من 4 إلى 5
2	من 5 إلى 6

(١٤) الأعداد التالية تمثل مراكز الفئات للتوزيع التكراري للعمليات التي تجرى يوميًا بإحدى المستشفيات .

3, 8, 13, 18, 23, 28, 33

(i) أوجد حدود هذه الفئات .

(ii) أوجد طول الفئة .

(١٥) الجدول التكراري التالي يبين درجات 180 طالبًا حصلوا عليها في أحد الاختبارات .

الدرجـــات	0 – 19	20 – 39	40 – 59	60 - 79	80 – 99
التكـــرار	18	51	66	32	13

(i) أوجد الحدود الحقيقية لهذه الفئات ومراكزها .

(ii) ارسم المدرج التكراري ، المضلع التكراري والمنحنى التكراري .

(iii) أوجد التكرار المتجمع الصاعد ومثله بيانيًا .

(١٦) البيانات التالية تمثل أطوال نوع من الزهور بالسنتيمترات .

4.1	5.0	4.8	4.3	4.2	5.3	4.2	3.6	4.2	4.4
4.5	3.2	4.0	3.8	3.8	5.3	4.5	4.6	4.0	5.2
5.2	4.4	4.7	4.1	4.6	4.9	4.1	5.8	4.2	4.2
4.8	4.1	5.6	4.5	5.1	4.6	4.3	5.2	4.7	3.2
4.0	4.6	4.0	4.2	4.5	3.5	4.7	4.9	3.9	4.8
3.7	5.4	4.9	4.6	4.3	5.4	5.0	4.5	4.7	4.3

(i) لخص البيانات أعلاه في جدول تكراري بالفئات التالية:

(3.0 – 3.4) , (3.5 – 3.9) , (4.0 – 4.4) , ... , (5.6 – 5.9)

(ii) ارسم المدرج التكراري النسبي.

(iii) أوجد التكرار النسبي المتجمع الصاعد ومثله بيانيًا.

(iv) كم عدد الزهرات التي تقل أطوالها عن 4.6 سم.

(١٧) ضع إشارة (،) إذا كان الجواب صحيحًا أو إشارة (X) إذا كان الجواب خطأ:

ا ـ المشاهدة «فصيلة الدم A» هي مشاهدة كمية ☐

ب ـ عندما تتساوى أطوال الفئات في جدول تكراري يكون الفرق بين مركزي فئتين متتاليتين مساويًا لطول الفئة ☐

جـ ـ العرض البياني يعطي فكرة أسرع وأوضح من العرض الجدولي للبيانات ☐

د ـ البيانات الوصفية تمثل بواسطة الخط البياني ☐

(١٨) أكمل ما يلي:

ا ـ يغلق المضلع التكراري المفتوح باقتراح فئتين ....

ب ـ البيانات الوصفية مثل: ....

جـ ـ البيانات الكمية مثل: ....

د ـ التمثيل بالأعمدة المزدوجة لظاهرتين هو: ....

هـ ـ التمثيل بالأعمدة المجزأة لظاهرتين هو: ....

و ـ المدرج التكراري هو عبارة عن: ....

ز ـ المنحنيات التكرارية غير المتماثلة لها نوعان مهمان هما: ....

ح ـ المنحنيات التكرارية المتماثلة هي عبارة عن: ....

ط ـ مجموع التكرارات النسبية في أي جدول توزيع تكراري يساوي: ......

٣

مقاييس النزعة المركزية أو الموضع
Averages and Measures of Central Tendency

● مقدمة ● الوسط الحسابي أو المتوسط ● الوسط المرجح أو الموزون ● الوسيط ● المنوال ● الوسط الهندسي ● الوسط التوافقي ● الربيعيات والعشيرات والمئينات ● تمارين

(٣ - ١) مقدمة

رأينا في الفصل السابق كيفية عرض البيانات الإحصائية وتلخيصها في جداول تكرارية أو رسوم بيانية بهدف الحصول على بعض الخصائص للمجتمع الإحصائي محل الدراسة. ومن المعروف عادة أن الرسوم البيانية تكون غير دقيقة لذلك يجب أن يكون لدينا مقاييس عددية تصف لنا هذه البيانات. وسوف نتعرض في هذا الفصل إلى نوع مهم من المقاييس الإحصائية وهو ما يسمى بمقاييس النزعة المركزية أو مقاييس الموضع أو المتوسطات. وهي مقاييس عددية تعين موقع التوزيع وهي مهمة في حالة المقارنة بين التوزيعات المختلفة بوجه عام. وتكون فائدتها أكثر في حالة التوزيعات المتشابهة في طبيعتها وشكلها ولكنها مختلفة في مواقعها. **فمثلاً**: عند دراسة عيّنة من البيانات الإحصائية التي تخص بعض الأسر من البادية حسب فئات الانفاق الاستهلاكي السنوي، وعيّنة من البيانات الإحصائية تخص بعض الأسر في الحضر، حسب فئات الانفاق الاستهلاكي أيضًا. فإن حساب المتوسط السنوي للانفاق لكل من البادية والحضر يُمكننا من المقارنة بينهما.

ويمكن تعريف المتوسطات بأنها القيمة النموذجية الممثلة لمجموعة من البيانات. وحيث إن القيمـة النمـوذجية تميل إلى الوقوع في المركز لذلك فإنه يمكن أن تسمى المتوسطات بمقاييس النزعةالمركزية. وسوف نتناول في هذا الفصل دراسة المتوسطات في صور مختلفة، كل منها له مميزاته وعيوبه وهذا يعتمد على طبيعة البيانات والهدف من استخدامها. وبعض المتوسطات الأكثر شيوعًا والتي سوف نتناولها بالشرح والتفصيل والأمثلة وكيفية حسـابهـا. وهي كل من الـوسط الحسابي (المتوسط) والوسط المرجح والوسيط والمنوال والوسط الهندسي والوسط التوافقي وذلك في حالة البيانات المباشرة والمبوَّبة.

(٣ - ١ - ١) تعريف رمز التجميع Σ

إذا كان لدينا مجموعة من المشاهدات $x_1 , x_2 , \ldots\ldots\ldots , x_n$ فإن حاصل جمع هذه المشاهدات يمكن التعبير عنه كمايلي:

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + x_n$$

وفي بعض الأحيان يكتب $\Sigma\ x$ وهذا معناه حاصل جمع قيم x وإذا كان لدينا مجموعة أخرى من المشاهدات $y_1 , y_2 , \ldots\ldots\ldots y_n$ وكذلك مقدار ثابت c فإنه يمكن إثبات (بسهولة) بعض العلاقات الآتية:

$$\text{(i)} \quad \sum_{i=1}^{n} (x_i \pm y_i) = \sum_{i=1}^{n} x_i \pm \sum_{i=1}^{n} y_i$$

$$\text{(ii)} \quad \sum_{i=1}^{n} c = nc$$

$$\text{(iii)} \quad \sum_{i=1}^{n} cx_i = c \sum_{i=1}^{n} x_i$$

وهذه العلاقات قد تكون مفيدة في إثبات بعض الخصائص لبعض المقاييس المختلفة.

(٣ ـ ٢) الوسط الحسابي أو المتوسط (Mean or Arithmetic Mean)

المتوسط أو الوسط الحسابي يعتبر من أهم مقاييس الموضع والأكثر استخدامًا في الإحصاء والحياة العملية إذ يستخدم عادة في الكثير من المقارنات بين الظواهر المختلفة. ولو أسندت قيمة المتوسط لكل مشاهدة فإن مجموع هذه القيم الجديدة يكون مساويًا لمجموع المشاهدات الأصلية [انظر مثال (٣ ـ ١)] ويعرف كالتالي:

إذا كان لدينا مجموعة من المشاهدات للمتغير X وهي $x_1, x_2, \ldots, x_n$ فإن الوسط الحسابي يساوي حاصل جمع المشاهدات أو البيانات مقسومًا على عددها ويرمز له بالرمز $\bar{x}$ (ويقرأ x bar) وعليه فإن

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \text{(٣ ـ ١)}$$

مثال (٣ ـ ١)

إذا كانت درجات 5 طلاب في إحدى المواد هي:

60, 72, 40, 80, 63

أحسب الوسط الحسابي لدرجات الطلاب.

الحــل

$$\sum_{i=1}^{5} x_i = 60 + 72 + 40 + 80 + 63$$
$$= 315$$

وبالتعويض في القانون (٣ ـ ١) نحصل على

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{5}(315)$$
$$= 63$$

الآن لو عوَّضنا بدل القراءة الأولى 60 بالمتوسط 63 وبالقراءة الثانية 72 بالمتوسط 63 وبالقراءة الثالثة 40 بالمتوسط 63 . . . الخ نجد

$$\sum_{i=1}^{5} x_i = 63 + 63 + 63 + 63 + 63 = 315$$

وذلك كما ذكر في الملاحظة السابقة في تعريف المتوسط.

(٣ ـ ٢ ـ ١) الوسط الحسابي للبيانات المبوبة (grouped data)

إذا كان لدينا عدد k من الفئات ذات المراكز $x_1, x_2, \ldots, x_k$ ولها تكرارات f_1, f_2 ......, f_k على الترتيب فإن الوسط الحسابي يعطى بالعلاقة التالية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + f_k x_k}{f_1 + f_2 + \ldots + f_k}$$

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{1}{n} \sum_{i=1}^{k} x_i f_i \qquad (٣ - ٢)$$

حيث $n = \sum_{i=1}^{k} f_i$ ونوضح ذلك بالمثال التالي.

مثال (٣ ـ ٢)

أحسب متوسط أعمار الطلاب $\bar{x}$ للبيانات التالية:

فئات الأعمار	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل

ولسهولة الحل نضع الجدول التالي:

الفئات	مراكز الفئات x	التكرار f	f x
5-6	5.5	2	11
7-8	7.5	5	37.5
9-10	9.5	8	76.0
11-12	11.5	4	46.0
13-14	13.5	1	13.5
المجموع Σ		20	184

بالتعويض في القانون (٣ ـ ٢) نحصل على

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i = \frac{1}{20}(184) = 9.2$$ سنه

(٣ ـ ٢ ـ ٢) بعض خصائص الوسط الحسابي

سوف نعرض بعض خصائص الوسط الحسابي وذلك في حالة البيانات المباشرة حيث يكون من السهل على الطالب استخدام نفس الطريقة لإثباتها في حالة البيانات المبوبة .

الخاصية الأولى

المجموع الجبري لانحرافات القيم من وسطها الحسابي يساوي صفرًا أي أنه إذا كانت مجموعة المشاهدات هي $x_1, x_2, \ldots, x_n$ وانحرافاتها عن وسطها الحسابي هي $d_1, d_2, \ldots, d_n$ حيث

$$d_i = x_i - \bar{x} \quad , \quad i = 1, 2, \ldots, n$$

فإن :

$$\sum_{i=1}^{n} (x_i - \bar{x}) = \sum_{i=1}^{n} d_i = 0$$

الإثبات

حيث إن

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

أي أن

$$n\bar{x} = \sum_{i=1}^{n} x_i$$

فإن

$$\sum_{i=1}^{n} (x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x}$$

$$= \sum_{i=1}^{n} x_i - n\bar{x} = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} X_i = 0$$

الخاصية الثانية

إذا كان للمتغير X المشاهدات $x_1 , x_2 , \dots , x_n$ فإذا أضفنا أو طرحنا من القيم الأصلية للمشاهدات مقدارًا ثابتًا b ، فإن الانحرافات $d_1, d_2, \dots , d_n$ حيث $d_i = x_i \pm b$ (وحيث $i = 1, 2, \dots , n$) تعطي المتوسط كالتالي :

$$\bar{x} = \bar{d} \mp b$$

الإثبات

حيث $d_i = x_i \pm b, i = 1, 2, \dots , n$

فإن :
$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} (x_i \pm b) = \sum_{i=1}^{n} x_i \pm nb$$

بالقسمة على n فإن
$$\frac{1}{n} \sum_{i=1}^{n} d_i = \frac{1}{n} \sum_{i=1}^{n} x_i \pm \frac{n}{n} b$$

أي أن
$$\bar{d} = \bar{x} \pm b$$

وعليه فإن
$$\bar{x} = \bar{d} \mp \bar{b}$$

وهو المطلوب .

مثال (٣ ـ ٣)

من مثال (٣ ـ ١) إذا أخذنا $b = 50$ فإن مجموع الانحرافات $\sum_{i=1}^{n} d_i$ يحسب كالتالي :

$$\sum_{i=1}^{5} d_i = (60-50) + (72-50) + (40-50) + (80-50) + (63-50)$$

$$\sum_{i=1}^{5} d_i = 10 + 22 - 10 + 30 + 13 = 65$$

$$\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i = \frac{1}{5}(65) = 13$$

ويكون
$$\bar{x} = \bar{d} + b = 13 + 50 = 63$$

وهي النتيجة نفسها التي حصلنا عليها في مثال (٣ ـ ١) .

الخاصية الثالثة

إذا كان للمتغير X المشاهدات $x_1, x_2, \ldots, x_n$ وضربنا هذه المشاهدات في مقدار ثابت حقيقي a فإن متوسط القيم الجديدة يساوي المتوسط $\bar{x}$ مضروبًا في a أي أن

$$\overline{(a\ x)} = a\,\bar{x}$$

ومن الخاصيتين الثانية والثالثة نجد أن :

$$\overline{(a\,x \pm b)} = a\,\bar{x} \pm b$$

الخاصية الرابعة

إذا كان للمتغير X المشاهدات $x_1, x_2, \ldots, x_n$ فإن مجموع مربعات انحرافات المشاهدات عن أي قيمة حقيقية c يكون أكبر أو يساوي مجموع مربعات انحرافات المشاهدات عن وسطها الحسابي أي أن :

$$\sum_{i=1}^{n} (x_i - c)^2 > \sum_{i=1}^{n} (x_i - \bar{x})^2$$

حيث $\bar{x} \neq c$.

وبرهان هذه الخاصية (أنظر الخاصية الثالثة للانحراف المعياري في الفصل الرابع).

(٣ ـ ٢ ـ ٣) بعض مميزات الوسط الحسابي

١) مقياس سهل حسابه ويخضع للعمليات الجبرية بسهولة.

٢) يأخذ في الاعتبار جميع القيم محل الدراسة.

٣) أكثر المقاييس فهمًا في الإحصاء.

(٣ ـ ٢ ـ ٤) بعض عيوب الوسط الحسابي

١) يتأثر بالقيم المتطرفة (وهي القيم الكبيرة جدًا أو الصغيرة جدًا مقارنة ببقية القيم).

٢) يصعب حسابه في حالة الجداول التكرارية المفتوحة حيث يتطلب معرفة مركز كل فئة. فعلى سبيل المثال: فتح فئة الأعمار الأولى مثال (٣-٢) من أسفل فتصبح (6 سنوات فأقل) بدلاً من (6–5) وفتح فئة الأعمار الأخيرة من أعلى لتصبح (13 سنة فأكثر) بدلاً من (14–13).

٣) يصعب حسابه في حالة البيانات الوصفية.

(٣ - ٣) الوسط المرجّح (Weighted mean)

في بعض الأحيان تكون المشاهدات $x_1, x_2, \ldots\ldots, x_n$ مقرونة بأوزان $w_1, w_2, \ldots\ldots, w_n$ في هذه الحالة فإن الوسط الحسابي يسمى بالوسط المرجح ويعطى بالعلاقة الآتية:

$$\bar{x}_w = \frac{w_1 x_1 + w_2 x_2 + \ldots + w_n x_n}{w_1 + w_2 + \ldots + w_n} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \qquad \ldots\ldots\ldots\ldots (٣ - ٣)$$

ونوضح ذلك بالمثال الآتي.

مثال (٣ - ٤)

أوجد الوسط المرجح $(\bar{x}_w)$ لدرجات طالب في ثلاث مواد إذا كانت الدرجات معطاة بالقيم 65 ,70 ,40 وكانت ساعات الدراسة الأسبوعية لهذه المواد بالترتيب هي:

2, 3, 4

الحل

$$\bar{x}_w = \frac{w_1x_1 + w_2x_2 + w_3x_3}{w_1 + w_2 + w_3}$$

فإن

$$x_1 = 40 \quad , \quad x_2 = 70 \quad , \quad x_3 = 65$$

$$w_1 = 2 \quad , \quad w_2 = 3 \quad , \quad w_3 = 4$$

وعليه فإن

$$\bar{x}_w = \frac{2 \times 40 + 3 \times 70 + 4 \times 65}{2 + 3 + 4}$$

$$= \frac{1}{9}(550)$$

$$= 61.11 \qquad \text{درجة}$$

ملاحظة

يمكن اعتبار الوسط الحسابي من الجداول التكرارية وسطًا مرجحًا حيث إنه يمكن وضع القيم $[w_1 = f_1 , w_2 = f_2 , \ldots\ldots , w_k = f_k]$ في القانون (٣ ـ ٣). ويكون

$$\bar{x}_w = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_k x_k}{f_1 + f_2 + \ldots + f_k} \quad \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$$

(٣ ـ ٤) الوسيط* Median

عند ترتيب البيانات (أو المشاهدات) ترتيبًا تصاعديًا (أو تنازليًا) فالوسيط يكون هو القيمة التي يقع 50% من البيانات قبلها في الترتيب و 50% من البيانات بعدها في الترتيب. فإذا كان عدد البيانات فرديًا يكون الوسيط هو المشاهدة التي في المنتصف وإذا كان عدد البيانات زوجيًا فإن الوسيط هو متوسط المشاهدتين اللتين في المنتصف.

مثال (٣ ـ ٥)

أوجد الوسيط لدرجات الطلاّب في مثال (٣ ـ ١) حيث كانت درجاتهم كالتالي:

60, 72, 40, 80, 63

---

* ويمكن إعطاء صيغة رياضية عامة للوسيط كالتالي:

إذا رمزنا بـ $P(X \leq x)$ إلى نسبة الوحدات الإحصائية التي لها قيمة x أو أقل فإن الوسيط هو Med بحيث:

$$P(X \leq Med) \geq 0.5 \quad , \quad P(X \geq Med) \geq 0.5$$

الحـــل

نرتب البيانات تصاعديًا:

$$40, 60, \boxed{63}, 72, 80$$

عدد المشاهدات فردي فيكون الوسيط المشاهدة التي في المنتصف.

الوسيط هو 63 (لاحظ أن بيانين يقعان في الترتيب قبل الوسيط وبيانين بعده).

مثال (٣ - ٦)

أوجد الوسيط لدرجات الطلّاب الآتية:

$$72, 60, 72, 40, 80, 63$$

الحـــل

نرتب البيانات تصاعديًا

$$40, 60, \boxed{63}, \boxed{72}, 72, 80$$

حيث إن عدد البيانات زوجي فيكون الوسيط هو متوسط المشاهدتين اللتين في المنتصف.

$$\text{Med} = \frac{63 + 72}{2} = 67.5$$

(٣ - ٤ - ١) الوسيط في حالة البيانات المبوبة (الجداول التكرارية)

قبل إيجاد قيمة الوسيط حسابيًا وبيانيًا تعرف الفئة الوسيطية كالتالي:

الفئة الوسيطية

هي الفئة التي يقع فيها الوسيط.

الوسيط حسابيًا

نتبع الخطوات التالية:

١) نكوّن الجدول المتجمع الصاعد (باستخدام الحدود الحقيقية).

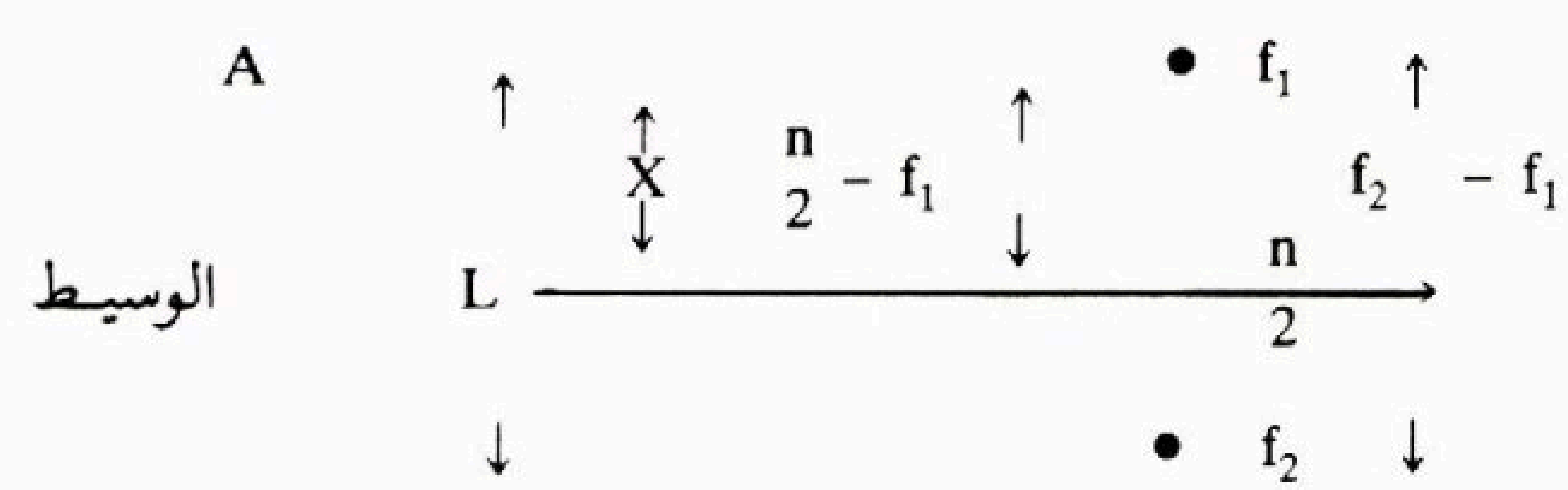
٢) نوجد رتبة الوسيط $\frac{n}{2}$ سواء أكانت n فردية أم زوجية .

٣) نحـدد مكـان الـوسيط بعد حساب $\frac{n}{2}$ بين التكرارات المتجمعة في الجدول المتجمع الصـاعـد ونضـع خطًا أفقيًا يمر داخل الفئة الوسيطية ويكون التكرار المتجمع السابق لهذا الخط هو f_1 والتكرار المتجمع اللاحق له هو f_2 ثم نحدد البداية الحقيقية للفئة الوسيطية ويرمز له بالرمز A ونعين طول الفئة الوسيطية ويساوي الحد الأدنى للفئة التالية مطروحًا منه الحد الأدنى للفئة الوسيطية ويرمز له بالرمز L ويعطى الوسيط بالعلاقة .

$$\text{Med} = A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} L \qquad \text{(٣ - ٤)}$$

حيث f_1 التكرار المتجمع الصاعد السابق للتكرار المتجمع الوسيطي f_2 التكرار المتجمع الصاعد اللاحق للتكرار المتجمع الوسيطي .

الإثبات



بإجراء التناسب بين الأطوال والتكرارات نحصل على :

$$\frac{x}{L} = \frac{\frac{n}{2} - f_1}{f_2 - f_1}$$

$$x = \frac{\frac{n}{2} - f_1}{f_2 - f_1} L$$

$$\text{Med} = A + x = A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} L$$

مثال (٣ ـ ٧)

أحسب الوسيط لأعمار الطلّاب في المثال (٣ ـ ٢) السابق.

الحـــل

نكوّن الجدول التكراري المتجمع الصاعد كالآتي:

التكرارات المتجمعة الصاعدة		فئات التكرارات المتجمعة الصاعدة		
	0	أقل من	4.5	
	2	أقل من	6.5	
f_1	7	أقل من	8.5	A
f_2	15	أقل من	10.5	
	19	أقل من	12.5	
	20	أقل من	14.5	

نحســـب $\frac{n}{2}$ وهـــي تســـاوي $\frac{20}{2} = 10$ ونـــلاحظ أن 10 تقـــع بين 15 ، 7 فنضع خطًّا أفقيًّا يمثل تكرار الوسيط المتجمع 10 وعليه فيكون

$$L = 10.5 - 8.5 = 2\ ,\ f_2 = 15\ ,\ f_1 = 7\ ,\ A = 8.5$$

وبتطبيق قانون الوسيط (٣ ـ ٤) نحصل على:

$$\text{Med} = 8.5 + \frac{10-7}{15-7}\ .2 = 8.5 + \frac{6}{8} = 9.25$$ (سنة)

الوسيط بيانيًّا

يمكن إيجـاد الـوسيط بيانيًّا من المنحنى المتجمع الصاعد أو المنحنى المتجمع الهـابط كل على حده أو تقاطع المنحنيين في رسم واحد، وفي حالة المنحنى المتجمع الصاعد تحدد نقطة $\frac{n}{2}$ على المحور الرأسي للتكرارات ونرسم منها خطًّا أفقيًّا موازيًا لمحور الفئات إلى أن يلتقي بالمنحنى في نقطة. نسقط من تلك النقطة عمودًا رأسيًّا يلاقي محور الفئات في نقطة تكون قيمتها هي قيمة الوسيط بيانيًّا. هذا ويمكن الحصول على قيمة

الوسيط أيضًا من المتجمع الهابط ، بأن نتّبع الخطوات السابقة نفسها والتي اتّبعناها في حالة المتجمع الصاعد لتحديد قيمة الوسيط بيانيًّا . أما في حالة تقاطع الصاعد والهابط في نقطة ، نسقط منها عمودًا رأسيًّا على محور الفئات تكون نقطة تقاطعه مع محور الفئات هي قيمة الوسيط ونوضح ذلك بالمثال التالي .

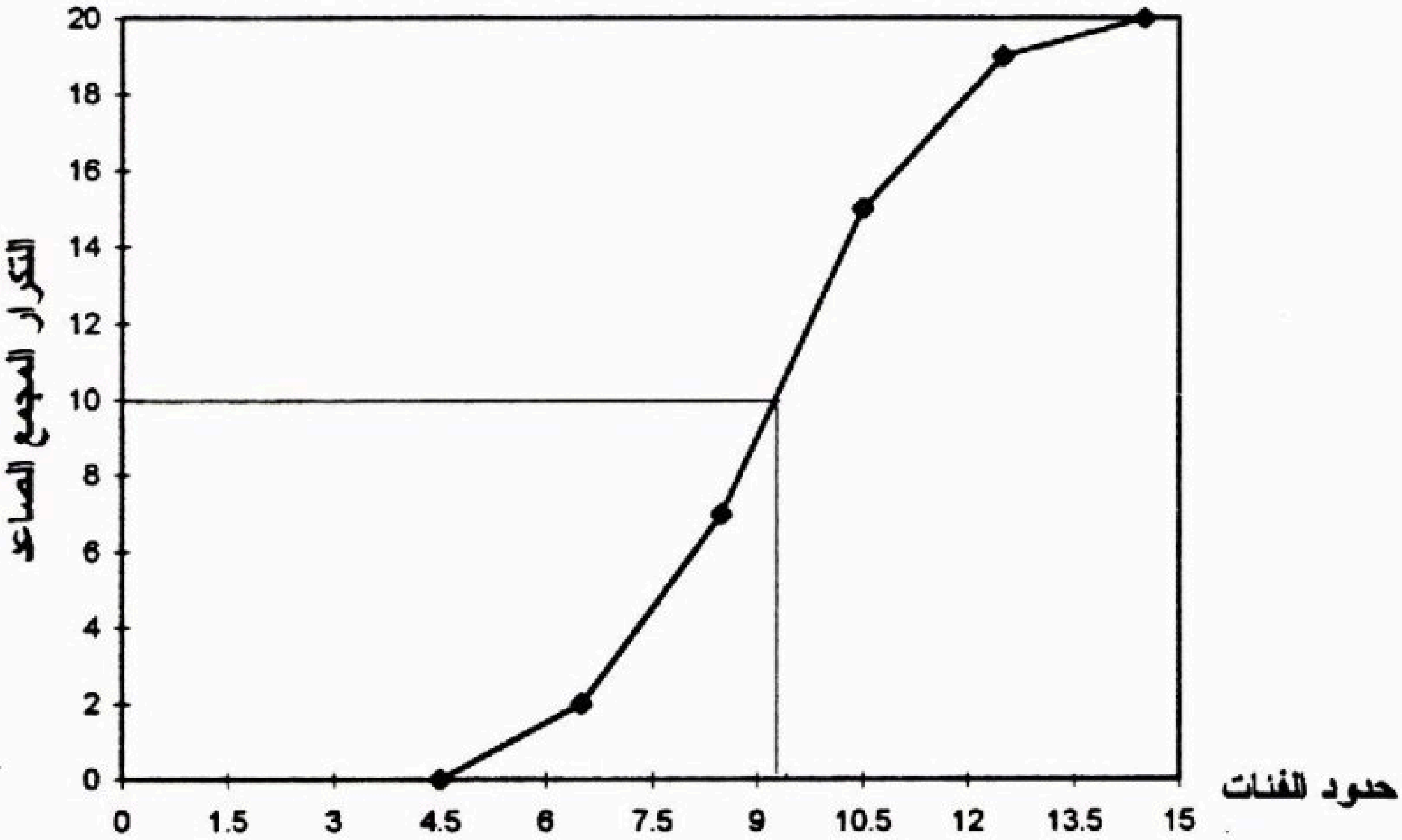
مثال (٣ ـ ٨)

أحسب الوسيط بيانيًّا من مثال (٣ ـ ٧) لأعمار الطلاب .

الحـــل

أولاً : الوسيط من المنحنى المتجمع الصاعد

من الجدول المتجمع الصاعد في مثال (٣ ـ ٧) يمكن رسم المنحنى المتجمع الصاعد شكل (٣ ـ ١) التالي :



شكل (٣ ـ ١) يمثل المنحنى المتجمع الصاعد لإيجاد قيمة الوسيط لأعمار الطلاب بيانيًّا

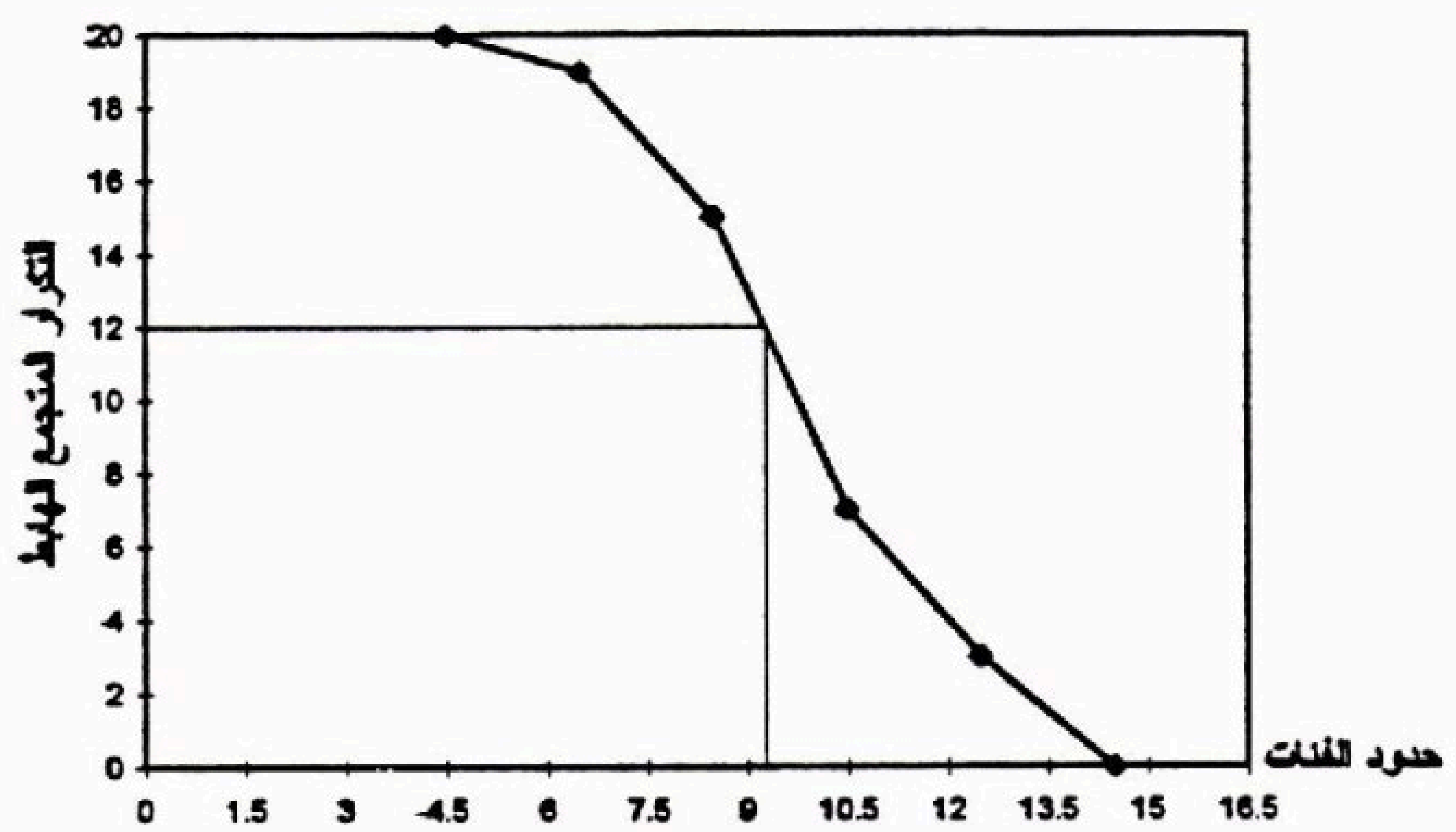
أي أن الوسيط هو القيمة التي عند النقطة C ويساوي تسع سنوات تقريبًا.

ثانيًا : الوسيط من المنحنى المتجمع الهابط

نكوّن الجدول المتجمع الهابط كالتالي :

التكرارات المتجمعة الهابطة	فئات التكرارات المتجمعة الهابطة
20	أكبر من أو يساوي 4.5
18	أكبر من أو يساوي 6.5
13	أكبر من أو يساوي 8.5
5	أكبر من أو يساوي 10.5
1	أكبر من أو يساوي 12.5
0	أكبر من أو يساوي 14.5

نرسم من هذا الجدول المنحنى المتجمع الهابط،شكل (٣ ـ ٢) التالي :

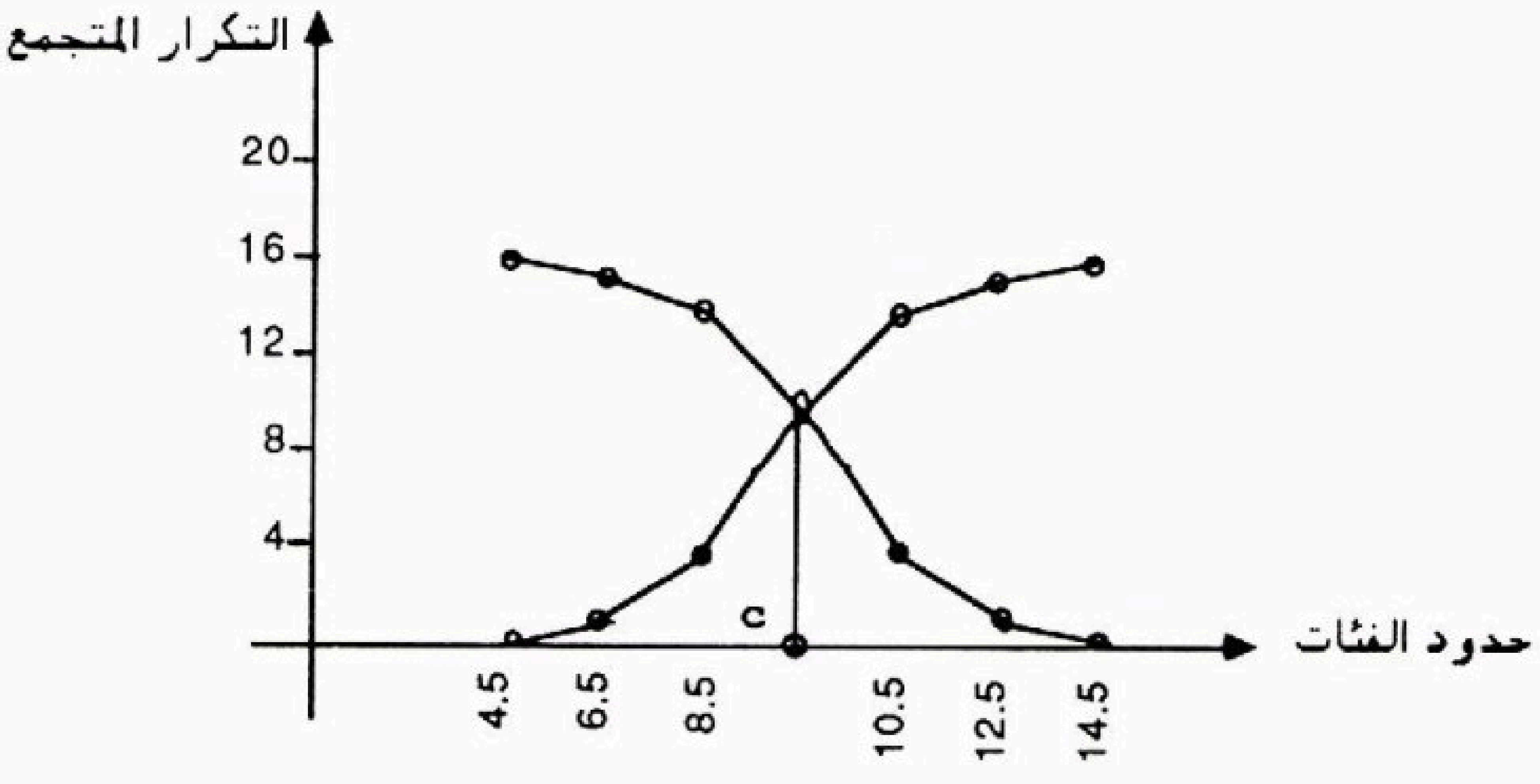


شكل (٣ ـ ٢) يمثل المنحنى المتجمع الهابط لإيجاد قيمة الوسيط لأعمار الطلاب بيانيًا

أي أن الوسيط هو القيمة عند النقطة C ويساوي تسع سنوات تقريبًا.

ثالثًا: الوسيط من المنحنيين المتجمعين الصاعد والهابط

نرسم المنحنين الصاعد والهابط من الجدولين الصاعد والهابط شكل (٣ ـ ٣) التالي:

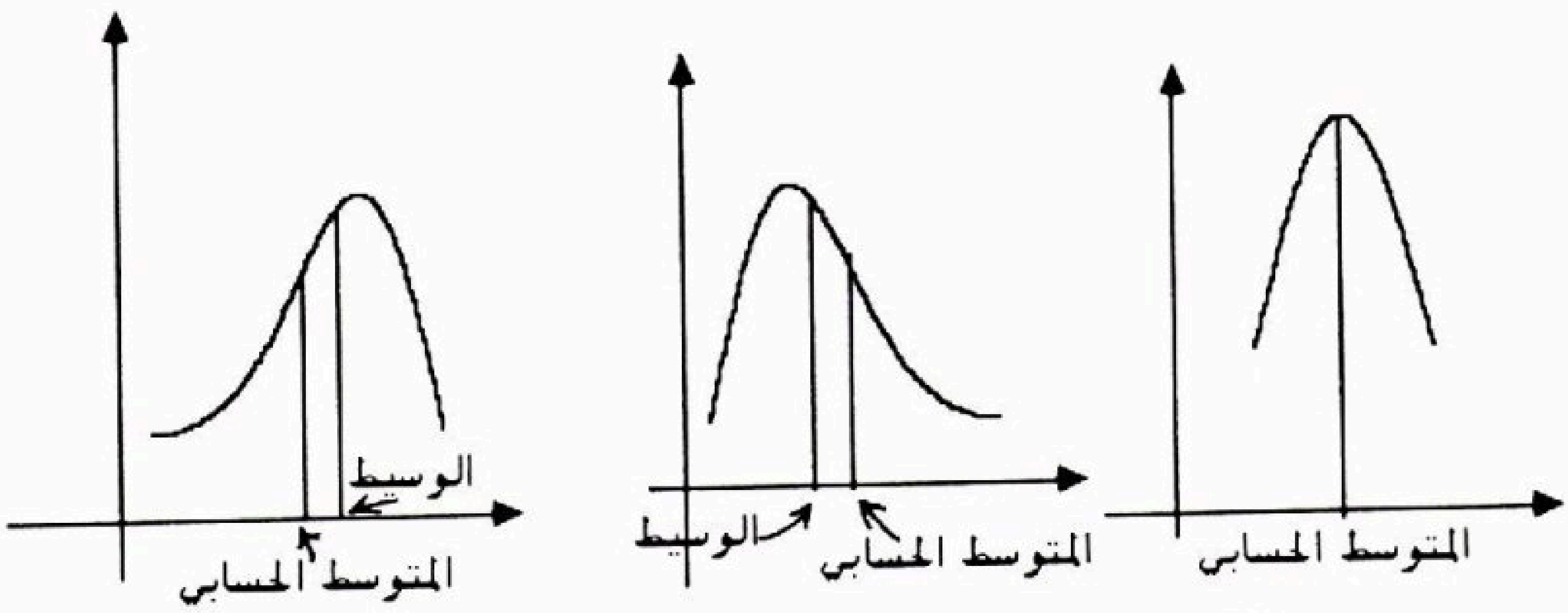


شكل (٣ ـ ٣) يمثل المنحنيين الصاعد والهابط لإيجاد الوسيط لأعمار الطلاب بيانيًا.

أي أن الوسيط هو القيمة عند النقطة C ويساوي تسع سنوات تقريبًا.

ملاحظة

يمكن الإشارة إلى أن الوسيط أكبر من الوسط الحسابي في حالة البيانات التي لها منحنيات تكرارية ملتوية نحو اليسار وأصغر من الوسط الحسابي في حالة البيانات التي لها تكرارات ملتوية نحو اليمين ويكون مساويًا للوسط الحسابي في المنحنيات المتماثلة. انظر الشكل (٣ ـ ٤) التالي:



شكل (٣ - ٤) يمثل وضع الوسيط بالنسبة للمتوسط الحسابي

(٣ - ٤ - ٢) مميزات الوسيط

١) لا يتأثر بالقيم المتطرفة للبيانات .

٢) يمكن حساب الوسيط في حالة الجداول التكرارية المفتوحة للبيانات الكمية .

٣) يمكن إيجاده في حالة البيانات الوصفية التي يمكن ترتيبها .

٤) مجموع الانحرافات المطلقة عن الوسيط أقل ما يمكن ، مقارنة بأي قيمة حقيقية a أي أن :

$$\sum \mid x_1 - \text{Med} \mid \leqslant \sum \mid x_i - a \mid$$

حيث $a \neq \text{Med}$

(٣ - ٤ - ٣) عيوب الوسيط

١) لا يأخذ جميع القيم في الاعتبار عند حسابه .

٢) لا يسهل التعامل معه في التحاليل الإحصائية والرياضية .

(٣ - ٥) المنوال (Mode)

هو القيمة الأكثر تكرارًا في مجموعة البيانات . وقد يكون لمجموعة البيانات منوال واحد ولذلك تسمى وحيدة المنوال أو يكون لها أكثر من منوال وتسمى متعددة المنوال . وقد لا يكون لمجموعة البيانات منوال بذلك تسمى عديمة المنوال .

مثال (٣ ـ ٩)

احسب المنوال من البيانات التالية :

2, 6, 9, 4, 6, 10, 6

الحـــل

يوجد لهذه البيانات منوال واحد وهي القيمة 6 لأنها تكررت 3 مرات أكثر من غيرها .

مثال (٣ ـ ١٠)

أحسب المنوال من البيانات التالية :

4. 2, 7, 9, 4, 7, 10, 7

الحـــل

نجد أن القيمة 7 تكررت 3 مرات والقيمة 4 تكررت مرتين وعليه فإن المنوال = 7 .

مثال (٣ ـ ١١)

أحسب المنوال من البيانات التالية :

4, 7, 4, 7, 8, 9, 7, 4, 10

الحـــل

نجد من البيانات بأن القيمة 4 تكررت 3 مرات والقيمة 7 تكررت 3 مرات أيضًا فإن هذه البيانات يوجد لها منوالان هما 7, 4 .

مثال (٣ ـ ١٢)

أحسب المنوال من البيانات التالية :

4, 9, 8, 12, 11. 7, 15

الحـــل

لا يوجد في هذه البيانات أي قيمة تكررت أكثر من مرة وعليه فإنه لا يوجد منوال لهذه البيانات .

(٣ - ٥ - ١) المنوال في حالة البيانات المبوبة (الجداول التكرارية)

في حالة البيانات المبوبة أو الجداول التكرارية لا يمكن القول بأن قيمة معيّنة يكون لها أكبر تكرار لأن القيم تذوب داخل الفئات المختلفة ولذلك يمكن القول بإنه توجد فئات منوالية وهي الفئة التي يقابلها أكبر تكرار وفي حالة تساوي تكرار الفئة السابقة للفئة المنوالية مع تكرار الفئة اللاحقة للفئة المنوالية فإنه يمكن حساب قيمة المنوال بمركز الفئة المنوالية أي في منتصفها وفي حالة عدم تساويهما في التكرار فإنه يمكن حساب المنوال بطريقة حسابية وبيانية كالآتي.

المنوال حسابيًا

نتبع الخطوات الآتية

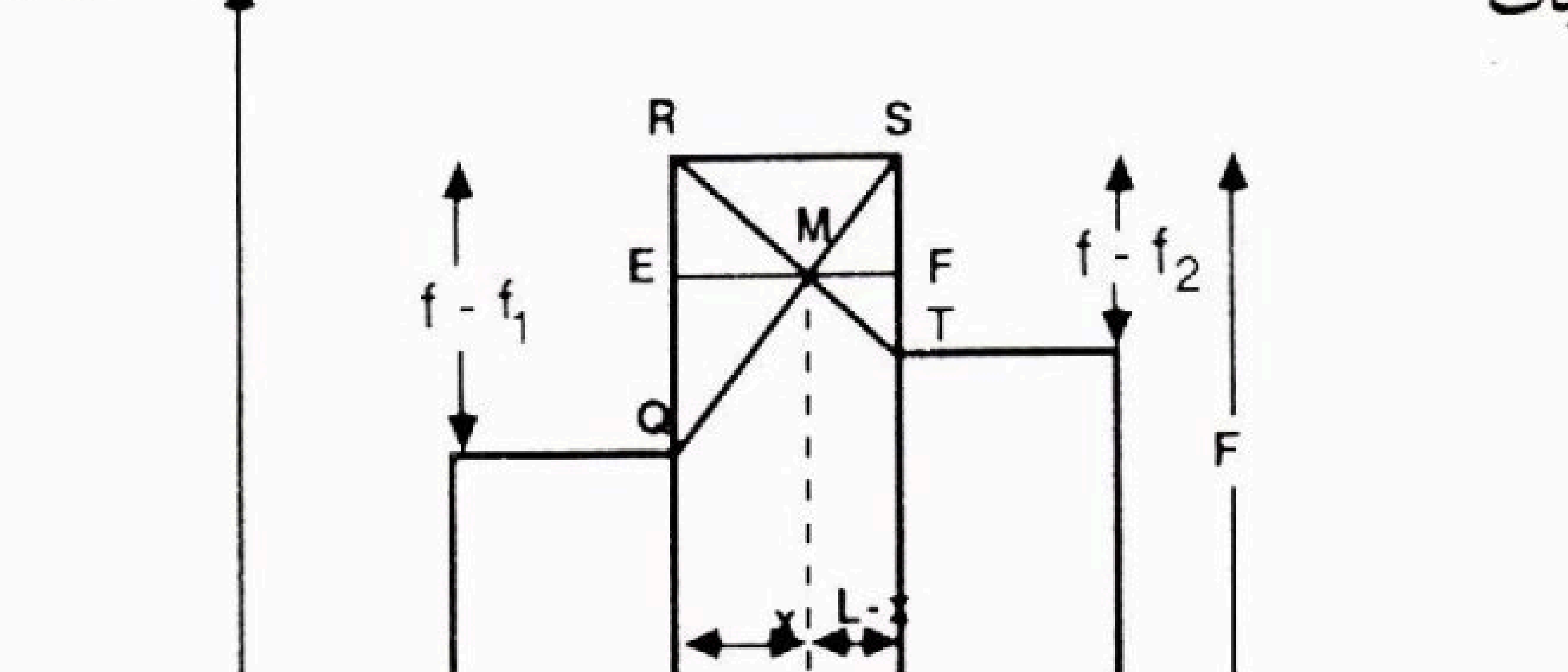
١) نوجد أكبر تكرار f وعليه يمكن إيجاد التكرار السابق له وهو f_1 والتكرار اللاحق له f_2 .

٢) نأخذ بداية الفئة المنوالية ويرمز له بالرمز A وهي الفئة التي تقابل أكبر تكرار f .

٣) نحدد طول الفئة المنوالية L وهو يساوي الفرق بين بداية الفئة المنوالية وبداية الفئة التالية لها ونطبق القانون الآتي:

$$Mod = A + \frac{f - f_1}{2f - f_1 - f_2} . L \quad \text{......} \quad (٣ - ٥)$$

الإثبات



شكل (٣ - ٥) يمثل التكرار المتوالي والسابق واللاحق له

من تشابه المثلثين MST, MQR في شكل (٣ ـ ٥) السابق نحصل على :

$$\frac{x}{f - f_1} = \frac{L - x}{f - f_2}$$

$$x = \frac{f - f_1}{2f - f_2 - f_1} \, . \, L$$

أي أن المنوال

$$\text{Mod} = A + x = A + \frac{f - f_1}{2f - f_2 - f_1} L$$

مثال (٣ ـ ١٣)

اوجد المنوال (حسابيًا) لأعمار الطلّاب في مثال (٣ ـ ٢).

الحـــل

أعمار الطلّاب كما يلي :

فئات الأعمار	5-6	7-8	9-10	11-12	13-14
التكـــرار	2	5	8	4	1

من الجدول نجد أن $f = 8$, $f_1 = 5$, $f_2 = 4$ وكذلك $A = 8.5$, $L = 10.5 - 8.5 = 2$

ونعوض في قانون المنوال السابق (٣ ـ ٥) نحصل على :

$$\text{Mod} = 8.5 + \frac{8 - 5}{16 - 5 - 4} \times 2$$

$$= 9.36 \text{ (سنـــة)}$$

المنوال بيانيًا

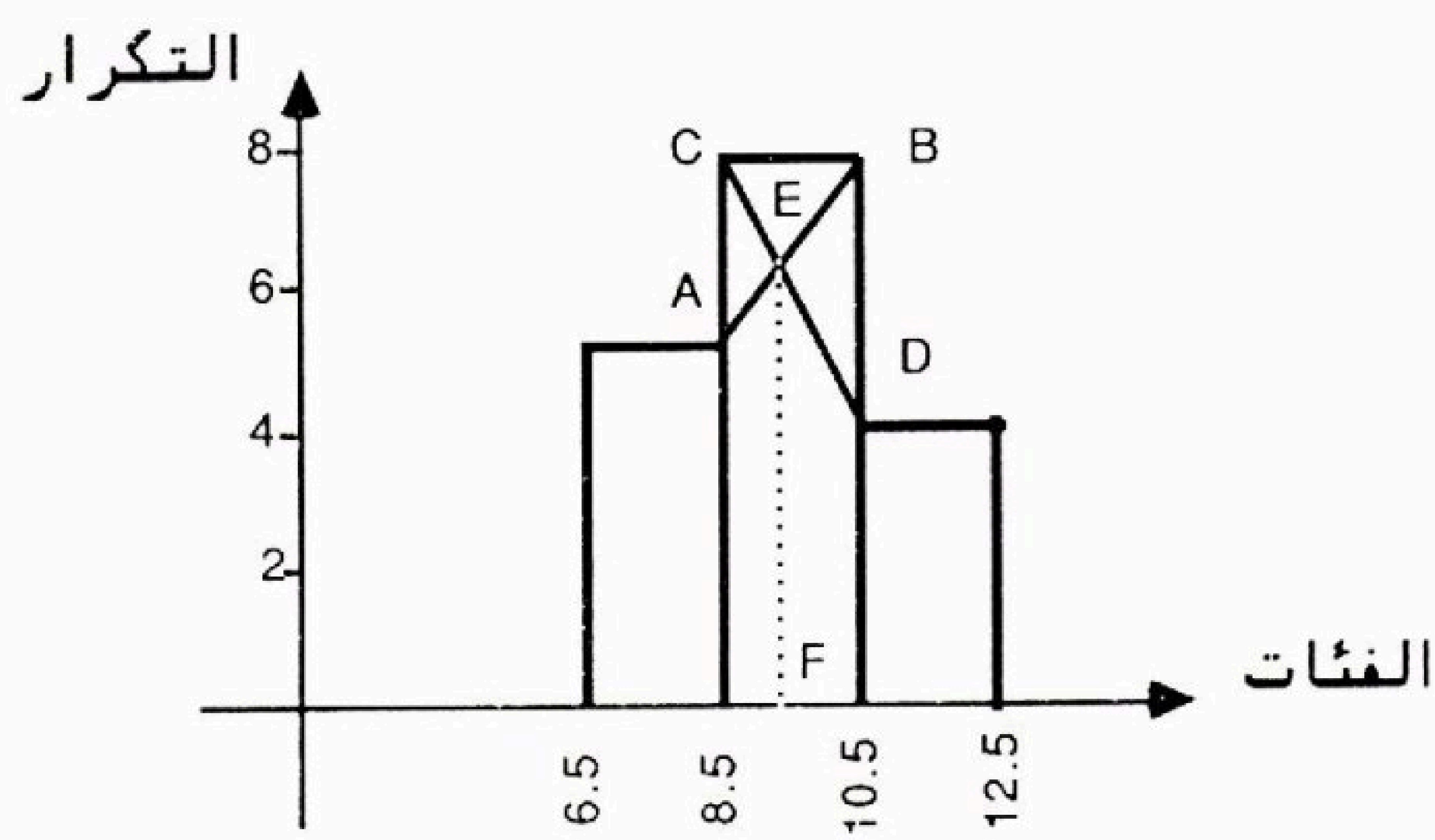
يكفي لإيجاد المنوال بيانيًا رسم ثلاث مستطيلات فقط من المدرج التكراري وهي المستطيل الممثل لأكبر تكرار والمستطيل السابق له وكذلك المستطيل اللاحق له ففي شكل (٣ ـ ٦) نصل النقطة A بالنقطة B ثم نصل النقطة C بالنقطة D فتكون نقطة E هي نقطة تقاطع المستقيمين CD و AB ويتحدد مكان المنوال بعد إسقاط عمود رأسي من E على محور الفئات الأفقي فتكون نقطة التقاطع F هي قيمة المنوال بيانيًا ونوضح ذلك بالمثال التالي:

مثال (٣ ـ ١٤)

احسب المنوال بيانيًا لأعمار الطلاب في مثال (٣ ـ ١٣).

الحــل

نرسم ثلاثة مستطيلات من المدرج التكراري كما في شكل (٣ ـ ٦) التالي:



شكل (٣ ـ ٦) يوضح قمية المنوال لأعمار الطلاب بيانيًا

أي أن المنوال القيمة عند النقطة F ويساوي تسع سنوات تقريبًا.

(٣ ـ ٥ ـ ٢) مميزات المنوال

١) مقياس سهل حسابه ولا يتأثر بالقيم المتطرفة.

٢) يمكن إيجاده للقيم الوصفية والتوزيعات التكرارية المفتوحة.

(٣ ـ ٥ ـ ٣) عيـوب المنـــوال

١) في حساب المنوال لا تؤخذ جميع قيم البيانات في الاعتبار.

٢) قد يكون لبعض البيانات أكثر من منوال وبذلك لا يمكن تحديد قيمة وحيدة للمنوال.

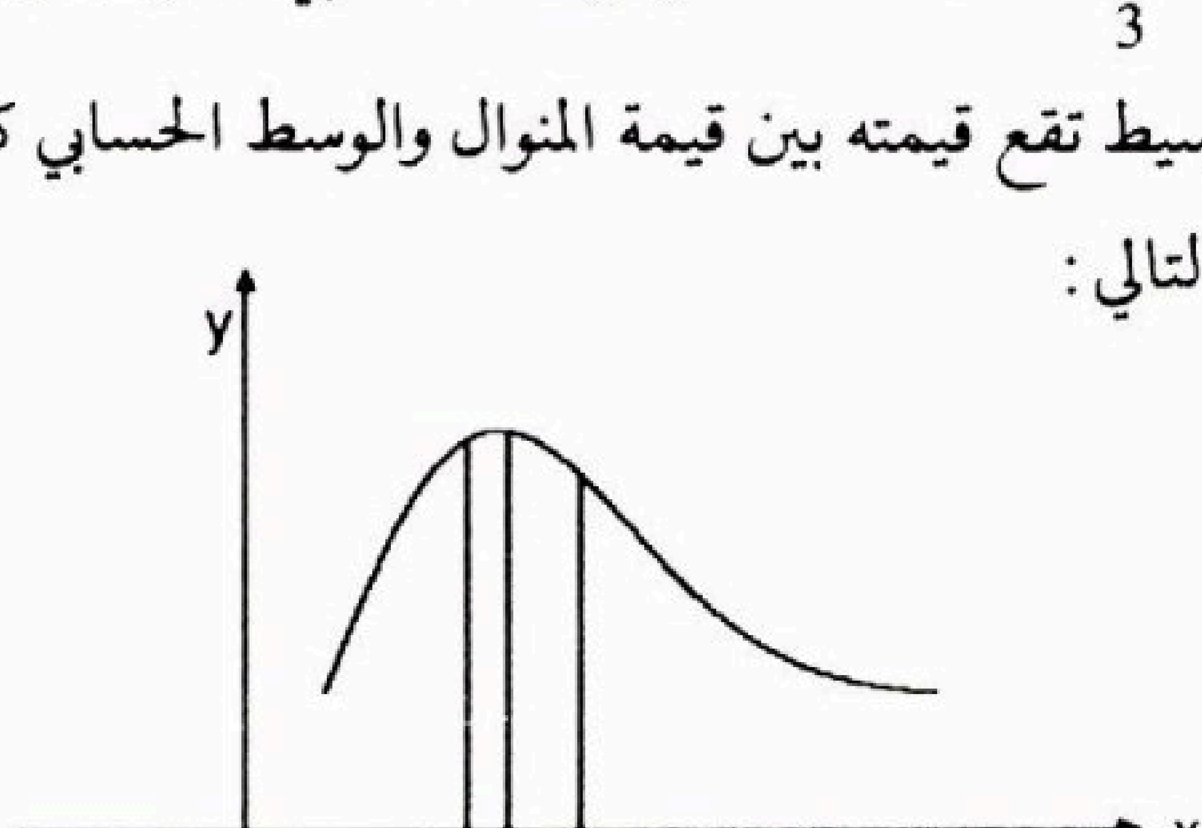
٣) في بعض الأحوال قد لا يوجد المنوال.

(٣ ـ ٦) العلاقة بين الوسط الحسابي والوسيط والمنوال

توجد علاقة تجريبية للمقاييس الثلاثة الوسط الحسابي، والوسيط والمنوال وذلك في حالة التوزيعات التكرارية أحادية المنوال وغير المتماثلة والمتماثلة وذات التواء بسيط وتعطى بالعلاقة الآتية:

$$\frac{\text{(الوسط الحسابي ـ المنوال)}}{3} = \text{(الوسط الحسابي ـ الوسيط)}$$

ولقد وجد أن الوسيط تقع قيمته بين قيمة المنوال والوسط الحسابي كما هو موضح بالرسم، شكل (٣ ـ ٧) التالي:



شكل (٣ ـ ٧) يوضح وضع الوسيط بالنسبة للمنوال والوسط الحسابي للمنحنيات ذات الالتواء البسيط

وفي حالة التوزيعات التكرارية المتماثلة الوحيدة المنوال تنطبق ثلاثة المقاييس وهي الوسط الحسابي، والوسيط، والمنوال فتكون قيمة واحدة وفي حالة التوزيعات التكرارية شديدة الالتواء فإن العلاقة السابقة تكون غير صحيحة ولا تصلح للاستخدام في تلك الحالة.

(٣ ـ ٧) الوسط الهندسي (Geometric mean)

الوسط الهندسي .G.M لمجموعة من القيم $x_1, x_2, \dots, x_n$ هو الجذر النوني لحاصل ضرب هذه القيم.

$$G.M = \sqrt[n]{x_1 \;\; x_2 \; \dots \; x_n} \quad \dots\dots\dots\dots \quad (٣ - ٦)$$

يمتاز الوسط الهندسي عن الوسط الحسابي بأنه أقل تأثرًا بالقيم المتطرفة في البيانات لأنه معلوم رياضيًا بأن الوسط الهندسي لمجموعة من القيم أقل من وسطها الحسابي وعادة يحسب الوسط الهندسي باستخدام اللوغاريتمات كالتالي:

$$\log G.M. = \frac{1}{n} \left(\sum_{i=1}^{n} \log x_i \right) \quad \dots\dots\dots\dots \quad (٣ - ٧)$$

ونوضح ذلك بالمثال الآتي.

مثال (٣ ـ ١٥)

احسب الوسط الهندسي والوسط الحسابي للبيانات

3, 5, 6, 6, 7, 10, 12

الحـــل

باستخدام العلاقة (٣ ـ ٦) نحصل على:

$$G.M. = \sqrt[7]{3.5.6.6.7.10.12}$$

وباستخدام اللوغاريتمات كما في العلاقة (٣ ـ ٧) يكون الوسط الهندسي

$$\log G.M. = \frac{1}{7} (\log 3 + \log 5 + \log 6 + \log 6 + \log 7 + \log 10 + \log 12)$$

$$= \frac{1}{7} (0.4771 + 0.6990 + 0.7782 + 0.7782 + 0.8451 + 1.0000 + 1.0729) = 0.8081$$

$$= 0.8081$$

وعليه فإن : $G.M. = 6.43$

والوسط الحسابي يكون $\bar{X} = \frac{1}{7}(3 + 5 + 6 + 6 + 7 + 10 + 12) = 7$

ونلاحظ أن قيمة الوسط الحسابي $\bar{X}$ من المثال أكبر من الوسط الهندسي .G.M وهذا يوضح حقيقة أن الوسط الهندسي لمجموعة أرقام موجبة غير متساوية، أقل من وسطها الحسابي .

(٣ ـ ٧ ـ ١) الوسط الهندسي في حالة البيانات المبوبة (الجداول التكرارية)

في هذه الحالة يحسب الوسط الهندسي للفئات التي عددها k ومراكزها هي $x_1, x_2, \ldots, x_k$ وهي التي يقابلها بالترتيب تكرارات $f_1, f_2, \ldots, f_k$ من القانون الآتي :

$$G.M = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \ldots \cdot x_k^{f_k}} \quad \ldots\ldots\ldots\ldots\ldots \quad (٣ ـ ٨)$$

حيث $n = \sum_{i=1}^{k} f_i$.

(٣ ـ ٨) الوسط التوافقي (Harmonic mean)

يستخدم الوسط التوافقي عندما يكون مقلوب المتغير له دلالة كأن يعين نسبة بين متغيرين مرتبطين مثل السرعة بالنسبة للزمن .

والوسط التوافقي H لمجموعة n من القيم $x_1, x_2, \ldots, x_n$ هو مقلوب الوسط الحسابي لمقلوبات هذه القيم أي أن :

$$H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} \quad \ldots\ldots\ldots \quad (٣ ـ ٩)$$

$$H = \frac{1}{(1/n)\sum_{i=1}^{n}(1/x_i)} = n \Big/ \sum_{i=1}^{n}(1/x_i)$$

ومن الناحية العملية يكون كالتالي :

$$\frac{1}{H} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{x_i} \qquad \text{.......................... (٣ - ١٠)}$$

مثال (٣ - ١٦)

أحسب الوسط التوافقي H للبيانات التالية :

3, 5, 6, 6, 7, 10, 12

الحـــل

باستخدام العلاقة (٣ - ١٠) نحصل على :

$$\frac{1}{H} = \frac{1}{7}\left(\frac{1}{3} + \frac{1}{5} + \frac{1}{6} + \frac{1}{6} + \frac{1}{7} + \frac{1}{10} + \frac{1}{12} \right.$$

$$\frac{1}{7}\left(\frac{140+84+70+70+60+42+35}{420} \right)$$

$$= \frac{501}{2940}$$

$$\therefore\ H = \frac{2940}{501} = 5.87$$

ويتضح من المثال السابق أن الوسط التوافقي (H = 5.87) أصغر من الوسط الهندسي (G.M. = 6.43) والوسط الحسابي ($\bar{x} = 7$) لنفس البيانات .

(٣ - ٨ - ١) الوسط التوافقي في حالة البيانات المبوبة (الجداول التكرارية)

في هذه الحالة إذا كان لدينا فئات عدد k ولها مراكز فئات : $x_1, x_2, \ldots , x_k$ ويقابلها بالترتيب التكرارات $f_1, f_2, \ldots , f_k$ فإن الوسط التوافقي في هذه الحالة يعطى بالعلاقة الآتية :

$$\frac{1}{H} = \frac{1}{n}\left(\frac{f_1}{x_1} + \frac{f_2}{x_2} + \ldots + \frac{f_k}{x_k}\right) = \frac{1}{n} \sum_{i=1}^{k} \frac{f_i}{x_i} \qquad \text{...... (٣ - ١١)}$$

وحيث إن $n = \sum_{i=1}^{k} f_i$ هي من مجموع التكرارات .

مثال (٣ - ١٧)

احسب الوسط الهندسي باستخدام العلاقة (٣ - ٨) ثم أحسب الوسط التوافقي باستخدام العلاقة (٣ - ١١) ثم قارن بينهما وذلك للجدول التكراري التالي الذي يمثل إنتاج 20 مزرعة بالطن لمحصول ما.

الفئات (الإنتاج بالطن)	1-3	4-6	7-9	10-12	13-15
التكرار (عدد المزارع)	2	3	8	5	2

الحل:

أولاً: الوسط الهندسي يمكن وضع العلاقة (٣ - ٨) في الصورة التالية:

$$\log G.M. = \frac{1}{n}\sum_{i=1}^{k} f_i \log x_i \qquad (٣-١٢)$$

ولسهولة الحل نضع الجدول التالي:

فئات الإنتاج	مركز الفئة x	التكرار f	f log x
1-3	2	2	2 log 2 = 0.206
4-6	5	3	3 log 5 = 2.097
7-9	8	8	8 log 8 = 7.225
10-12	11	5	5 log 11 = 5.207
13-15	14	2	2 log 14 = 2.292
Σ		20	17.027

بالتعويض في العلاقة (٣ - ١٢) السابقة من جدول الحل:

$$\log G.M. = \frac{1}{20}(17.027) = 0.85135$$

أي أن الوسط الهندسي G.M هو:

$$G.M = 6.924 \quad \text{طناً}$$

ثانيًا: الوسط التوافقي باستخدام العلاقة (٣ - ١١) ولسهولة الحل نضع الجدول التالي:

فئات الإنتاج	مركز الفئة x	التكرار f	$\frac{1}{x}$	$\frac{f}{x}$
1-3	2	2	0.5	1.00
4-6	5	3	0.33	0.99
7-9	8	8	0.125	1.00
10-12	11	5	0.2	1.00
13-15	14	2	0.5	1.00
		20		4.99

بالتعويض في العلاقة (٣ ـ ١١) من جدول الحل نجد أن:

$$\frac{1}{H} = \frac{1}{20}\sum_{i=1}^{5} f_i / x_i = \frac{1}{20}(4.92) = 0.2495$$

أي أن الوسط التوافقي H هو:

$$H = 1/0.2495 = 4.008 \quad \text{طناً}$$

يتضح من المثال السابق أن الوسط التوافقي $(H = 4.008)$ أصغر من الوسط الهندسي $(G.M. = 6.924)$.

(٣ - ٩) الربيعات والعشيرات والمئينات

إذا رتّبت عيّنة من البيانات حسب قيمها تصاعديًا أو تنازليًا فإن القراءة التي تكون في المنتصف والتي تقسم العيّنة إلى مجموعتين متساويتين في العدد هي الوسيط كما سبق تعريفه. وبتعميم الفكرة وتقسيم البيانات بعد ترتيبها إلى أربعة أجزاء متساوية فإن نقاط التقسيم يرمز لها بـ Q_1, Q_2, Q_3 حيث Q_1 يسمى الربيع الأول و Q_2 يسمى الربيع الثاني (تساوي الوسيط) و Q_3 يسمى الربيع الثالث.

وكذلك يمكن إيجاد القيم التي تقسم البيانات السابقة بعد ترتيبها إلى عشرة أقسام ونرمز لنقاط التقسيم بـ $D_1 D_2, \ldots ,D_{10}$ حيث D_1 يسمى العُشَيْر الأول وهو يمثل القيمة التي يسبقها عُشْر القراءات و D_2 تسمى العُشَيْر الثاني وهو يمثل القيمة التي يسبقها $(\frac{2}{10})$ من القراءات وهكذا. يمكن إيجاد القيم التي تقسم البيانات السابقة بعد ترتيبها إلى مائة قسم ونرمز لنقاط التقسيم $P_1, P_2, \ldots , P_{99}$ حيث P_1 يسمى المئين الأول وهو يمثل القيمة التي يسبقها $\frac{1}{100}$ من القراءات و P_2 يسمى المئين الثاني وهو يمثل القيمة التي يسبقها $\frac{2}{100}$ من القراءات وهكذا لباقي المئينات ويعطى قانون حساب العشيرات والمئينات في حالة البيانات المبوبة مثل قانون الوسيط السابق (٣ ـ ٤) مع استبدال $\frac{n}{2}$ بـ $\frac{n}{10}$ للعشير الأول، $\frac{2n}{10}$ للعشير الثاني وهكذا... واستبدال $\frac{n}{2}$ بـ $\frac{n}{100}$ للمئين الأول، $\frac{2n}{100}$ للمئين الثاني وهكذا... ونوضح ذلك بالمثال التالي.

مثال (٣ ـ ١٨)

احسب لأعمار الطلاب في مثال (٣ ـ ٢) السابق كل من العشير الثاني والمئين التسعين حسابيًا وبيانيًا.

الحـــل

المطلوب حسابيًّا

الجدول المتجمع الصاعد لأعمار الطلاب كالتالي:

	حـدود الفئات	التكرار المتجمع الصاعد	
	أقـــل مـــن 4.5	0	
	أقـــل مـــن 6.5	2	f_1
D_2	أقـــل مـــن 8.5	7	f_2
	أقـــل مـــن 10.5	15	f'_1
P_{90}	أقـــل مـــن 12.5	19	f'_2
	أقـــل مـــن 14.5	20	

لإيجاد العشير الثاني D_2 حسابيًّا نستخدم القانون

$$D_2 = A + \frac{\frac{2n}{10} - f_1}{f_2 - f_1} L$$

حيث A بداية الفئة للعشير الثاني. n مجموع التكرارات f_1 التكرار المتجمع السابق، f_2 التكرار المتجمع اللاحق، L طول فئة العشير الثاني.

$$D_2 = 6.5 + \frac{4-2}{7-2} \,.\, 2 = 6.5 + \frac{4}{5} = 7.3$$

لإيجاد المئين التسعين P_{90} حسابيًّا نستخدم القانون:

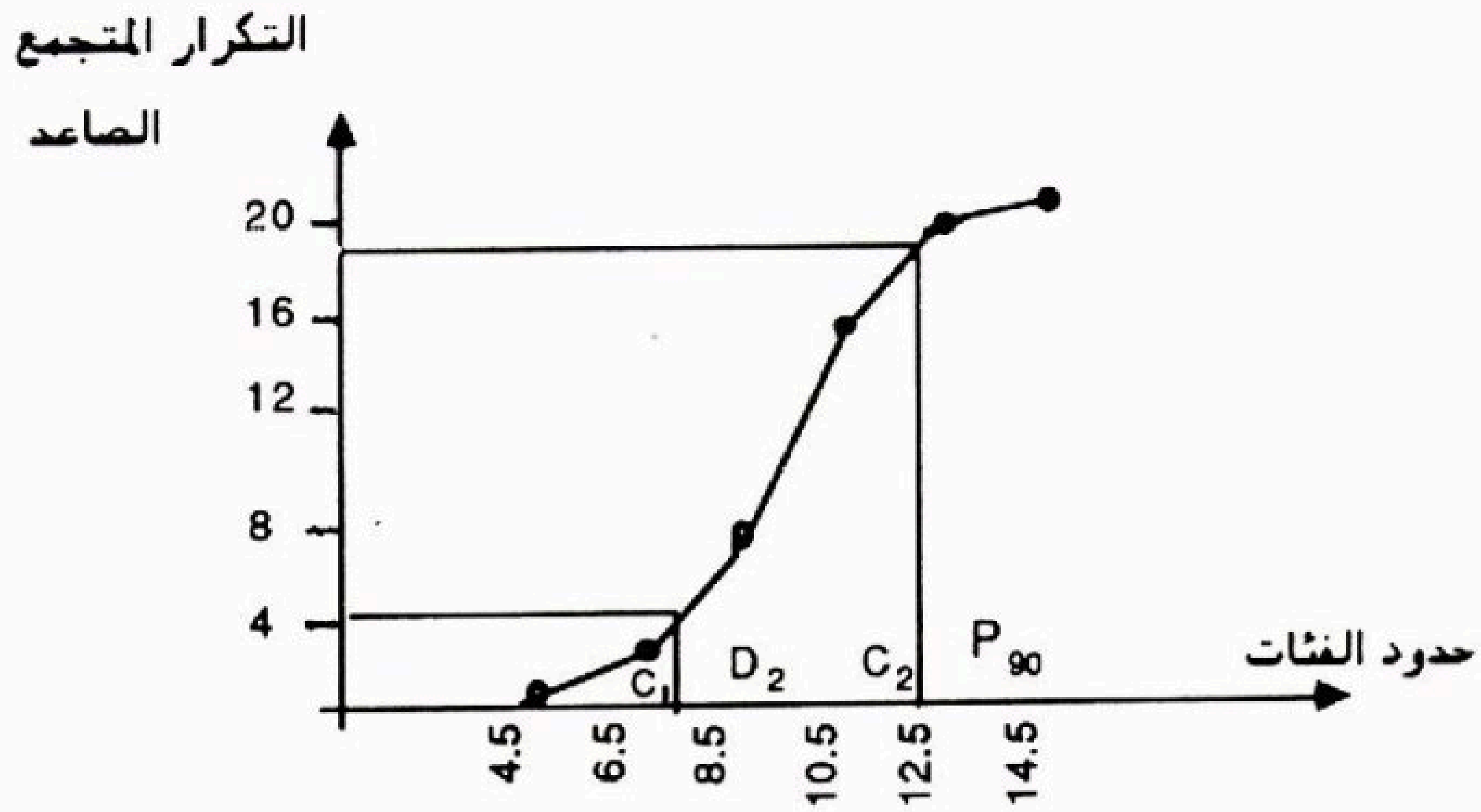
$$P_{90} = A' + \frac{\frac{90n}{100} - f'}{f'_2 - f'_1} L$$

$$= 10.5 + \frac{18-15}{19-15} 2 = 10.5 + \frac{6}{4}$$

$$= 12$$

المطلوب بيانيًا

نرسم المنحنى المتجمع الصاعد من الجدول المتجمع الصاعد كما في شكل (٣ - ٨) التالي :



شكل (٣ - ٨) يبين إيجاد قيمة كل من العشير الثاني والمئين التسعين بيانيًا

أي أن قيمة العشير الثاني (D_2) عند النقطة C_1 يساوي 7 تقريبًا وكذلك قيمة المئين التسعين (P_{90}) عند النقطة C_2 يساوي 12 تقريبًا .

(٣ - ١٠) تمارين

(١) المتغيران x, y كان لهما القيم الآتية :

X	1	2	4	−2	5
Y	−1	1	5	−7	7

أحسب قيمة كل مما يأتي :

(i) $\sum x, \sum y, \sum xy, \sum x^2, \sum y^2, \sum y^3, \sum x^3, \sum xy^2,$

(ii) $\sum(x+2)(x-2)\ ,\ \sum x(x+4)\ ,\ \sum(x+3)^2\ ,\ \sum(x-3y)^2\ ,\ (\sum x)^2\ ,\ ((\sum x)+2)^2$

(٢) إذا كانت c مقدارًا ثابتًا وكان المتغير x يأخذ القيم $x_1, x_2, \dots, x_n$ والمتغير y يأخذ القيم $y_1, y_2, \dots, y_n$ فاثبت أن :

(i) $\frac{1}{n} \sum (x+y) = \frac{1}{n} \sum x + \frac{1}{n} \sum y$

(ii) $\sum (cx+y) = c \sum x + \sum y$

(iii) $\sum \left(x - \frac{\sum x}{n}\right)^2 = \sum x^2 - \frac{(\sum x)^2}{n}$

(٣) أي من العلاقات الآتية صحيحة :

(i) $\sum \left(\frac{y}{x}\right) = \frac{\sum y}{\sum x}$

(ii) $\sum xy = \sum x \sum y$

(iii) $\sum (x-c)(x+c) = \sum x^2 - nc^2$

(iv) $\sum (x+y)^2 = \sum x^2 + \sum y^2 + 2\sum xy$

(v) $(\sum x)^2 = \sum x^2$

(٤) فيما يلي أعمار مجموعة من الطلاب بإحدى المدارس الابتدائية

6, 6, 9, 8, 6, 10, 9, 9, 8, 7, 8, 6, 7, 8, 8, 11, 10, 11, 8, 8,

ا) أحسب المتوسط الحسابي لأعمار هؤلاء الطلاب .

ب) أوجد المنوال لأعمار هؤلاء الطلاب .

جـ) أوجد الوسيط لأعمار هؤلاء الطلاب .

د) ما قيمة المقاييس الثلاثة بعد 3 سنوات بفرض بقائهم جميعًا على قيد الحياة .

(٥) فيما يأتي درجات أحد الطلاب في 5 امتحانات .

90, 40, 81, 72, 66

ا) أوجد الوسط الحسابي لهذه الامتحانات .

ب) إذا أضفنا درجتين لكل امتحان ما هو الوسط الحسابي للدرجات الجديدة؟

جـ) إذا ضربنا نتيجة كل امتحان في 2 ما هو الوسط الحسابي للدرجات الجديدة؟

(٦) عند فحص مجموعة من الأرقام يتكون كل منها من رقم واحد كانت البيانات على الصورة التالية :

التكرار	الرقم
8	2
10	3
20	5
20	7
6	8
6	9

أحسب الوسط الحسابي والوسيط والمنوال لهذه البيانات .

(٧) فيما يلي توزيع درجات 60 طالبًا في أحد الاختبارات :

فئات الدرجات	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99
عدد الطلاب	3	3	4	6	6	11	9	8	2	4	3	1

احسب الوسط الحسابي والوسيط والمنوال لدرجات الطلاب .

(٨) فيما يلي أطوال مجموعة من الطلاب في إحدى المدارس :

أطوال الطلاب	65	66	67	68	69	70	71	72	73	74
عدد الطلاب	2	0	2	0	3	6	4	3	4	1

أحسب الوسط الحسابي للأطوال وكذلك الوسيط والمنوال .

(٩) فيما يلي توزيع الأجر اليومي لعدد من العمال بالريال في أحد المصانع :

فئات الأجر	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79
عدد العمال	9	12	15	8	4	2

ا) أحسب الوسط الحسابي لأجورالعمال .

ب) أوجد الوسيط والمنوال حسابيًا وبيانيًا .

جـ) إذا كان الأجر اليومي لكل عامل يزيد بمقدار خمسة ريالات كل ستة شهور فما قيمة المقاييس السبقة بعد السنة؟

د) أوجد الوسط الهندسي والوسط التوافقي للأجور.

(١٠) أوجد الوسط الحسابي والوسط الهندسي والتوافقي لمجموعة الأرقام :

0, 2, 4, 6

(١١) إذا كانت لدينا البيانات التالية :

2, 10, 15, 8, 6, 17, 2, 10, 3, 9, 5, 9, 1, 10, 13

ا) أوجد المتوسط الحسابي والوسيط والمنوال .

ب) لو أضفنا إلى كل قراءة من القراءات أعلاه مقدارًا يساوي 3 فما التغير الذي يطرأ على مقاييس النزعة المركزية؟

جـ) لو ضربنا كل قراءة من البيانات أعلاه بالرقم 5 ، فماذا يطرأ من تغيير على البيانات أعلاه؟

(١٢) إذا كانت الحمولة القصوى لمصعد الكلية هي 3000 رطلاً فهل تعد الحمولات التالية أكبر من طاقة هذا المصعد.

ا) إذا صعد 13 طالبًا وزن كل منهم 165 رطلاً.

ب) إذا صعد 12 طالبًا وزن كل منهم 123 رطلاً وتسعة آخرين وزن كل منهم 175 رطلاً.

(١٣) إذا كانت أسعار أربعة أنواع من الفاكهة 26.5, 19.6, 36.9, 41.0 ريالاً على التوالي للصندوق، إذا باع تاجر ما 59 صندوقًا من النوع الأول، 156 صندوقًا من النوع الثاني، 386 صندوقًا من النوع الثالث، 8 صناديق من النوع الرابع:

أوجد متوسط سعر البيع للصندوق الواحد.

(١٤) الجدول التالي يبين متوسط دخل العمال في شركة ما حسب مهنة كل منهم:

نــــوع العمــــال	عــــدد العمــــال	متوسط الدخل الأسبوعي بالريال
عمال التصنيع	99,900	204.71
عمال المناجم	32,600	285.48
عمال التشييد	40,400	330.22

أوجد متوسط الدخل الأسبوعي للعمال الذين يعملون بهذه الشركة.

(١٥) فيما يلي تصنيف لعدد أيام الغياب خلال فصل دراسي لشعبة تتضمن 46 طالبًا.

4	3	2	1	0	عدد أيام الغياب x
3	5	8	10	20	عدد الطــلاب f_i

- احسب المنوال والمتوسط والوسيط لعدد أيام الغياب في هذه الشعبة.

٤

مقاييس التشتت
Measures of Dispersion

● مقدمة ● المدى ● نصف المدى الربيعي
● الانحراف المتوسط ● الانحراف المعياري
● معامل الاختلاف ● نظريف شيبشيف
● المتغير المعياري والدرجات المعيارية (مقياس
التمركز) ● مقاييس الالتواء (الشكل)
● التفرطح ● تمارين

(٤ - ١) مقدمة

لقد سبق لنا دراسة طرق عرض البيانات جدوليًا وبيانيًا والتعرف على أشكالها وتوزيعاتها المختلفة وكذلك دراسة مقاييس النزعة المركزية (المتوسطات) وذلك لوصف البيانات عدديًا لهذه التوزيعات المختلفة. ولكن طرق عرض البيانات وحساب المتوسطات للمجموعات المختلفة من البيانات غير كافٍ للمقارنة بين هذه المجموعات، ولتوضيح ذلك نأتي بمثال لدراسة درجات ثلاث مجموعات مختلفة من الطلاب z و y و x وكانت الدرجات كالآتي:

x : 59, 61, 62, 58, 60

y : 50, 60, 66, 54, 70

z : 19, 65, 46, 78, 72

وبحساب المتوسط الحسابي لثلاث المجموعات نجده يساوي 60 درجة لكل منها. ولكن عند النظر إلى درجات المجموعة الأولى نجدها متقاربة، ودرجات المجموعة

الثانية أقل تقاربًا من المجموعة الأولى، والمجموعة الثالثة أقل تقاربًا من المجموعة الثانية. أي أن ثلاث المجموعات مختلفة التجانس على الرغم من أن المتوسط الحسابي للثلاث المجموعات هو نفسه. وبذلك تكون مقاييس النزعة المركزية غير كافية للمقارنة بين طبيعة البيانات الإحصائية. لذلك نشأت الحاجة إلى إيجاد مقاييس تقيس درجة تجانس (تقارب) أو تشتت (تباعد) مفردات البيانات بعضها عن البعض وهذه المقاييس هي ما تسمى بمقاييس التشتت. وهي كثيرة وسوف نعرض منها: المدى، نصف المدى الربيعي، الانحراف المتوسط، التباين والانحراف المعياري، ومعامل الاختلاف (التغير) ومقاييس الالتواء والتفلطح، وسوف نتناول كلًا منها بالشرح والتفصيل والأمثلة كلًا على حده.

(٤ - ٢) المدى (Range)

يعرف المدى بأنه الفرق بين أكبر قراءة وأصغر قراءة في مجموعة قراءات. أي أن (المدى = أكبر قراءة - أصغر قراءة)، وذلك في حالة البيانات المباشرة. أما في حالة البيانات المبوبة فإن المدى يعرف بأكثر من طريقة نذكر منها فيما يلي طريقتين:

i) المدى = الفرق بين مركزي الفئة العليا والفئة الدنيا.

ii) المدى = الحد الأعلى للفئة العليا مطروحًا منه الحد الأدنى للفئة الدنيا

ونوضح ذلك بالأمثلة في حالة البيانات المباشرة والمبوبة كما يلي.

مثال (٤ - ١)

أحسب المدى (R) لدرجات الطلاب الآتية:

82, 40, 62, 70, 30, 80

الحل

أكبر قراءة = 82 درجة، أصغر قراءة = 30 درجة.

$R = 82 - 30 = 52$ (درجة)

مثال (٤ ـ ٢)

أوجد المدى (R) لدرجات مجموعة من الطلاب معطاة بالجدول الآتي :

الفئـــات	40 – 49	50 – 59	60 – 69	70 – 79	80 – 89	90 – 99
التكـــرار	2	9	15	11	2	1

الحـــل

يمكن إيجاد المدى بطريقتين :

الطريقة الأولى

مركز الفئة العليا = 94.5 ومركز الفئة الدنيا = 44.5

$R = 94.5 - 44.5 = 50$ (درجة)

الطريقة الثانية

الحد الأعلى للفئة العليا (الحقيقي) = 99.9

الحد الأدنى للفئة الدنيا (الحقيقي) = 39.5

$R = 99.5 - 39.5 = 60$ (درجة)

ونلاحظ اختلاف كل من الطريقتين في حساب قيمة المدى وتستخدم الطريقة الأولى غالبًا في إيجاد المدى.

(٤ ـ ٢ ـ ١) بعض مميزات المدى

١) سهل الحساب جدًّا.

٢) يعطي فكرة سريعة عن طبيعة البيانات ويستخدم كثيراً في مراقبة جودة الإنتاج وكذلك في وصف الأحوال الجوية.

(٤ ـ ٢ ـ ٢) بعض عيوب المدى

١) يعتمد في حسابه على قيمتين فقط من البيانات مع إهمال باقي البيانات.

٢) يتأثر كثيراً بالقيم المتطرفة (الشاذة) لذلك فهو مقياس تقريبي لا يعتمد عليه.

(٤ - ٣) نصف المدى الربيعي (Semi-interquartile range)

من أهم عيوب المدى أنه يتأثر كثيراً بالقيم الشاذة ومن ثم فهو لا يعطي صورة صادقة عن طبيعة البيانات ؛ لذلك دعت الحاجة إلى إيجاد مقياس آخر يتخلص من تأثير هذه القيم الشاذة وهذا المقياس هو ما يسمى **بنصف المدى الربيعي** ويعرف كما يلي :

إذا كانت لدينا مجموعة من البيانات عددها n قراءة فإن القراءات ترتب ترتيبًا تصاعديًا وتقسم إلى أربعة أقسام متساوية كما هو موضح على الخط الأفقي التالي :

$Q_1 \qquad Q_2 \qquad Q_3$

تسمى القيمة التي يسبقها ربع القراءات بالربيع الأدنى ويرمز لها بالرمز Q_1 ورتبته $\frac{n}{4}$ وتسمى القيمة التي يسبقها ثلاثة أرباع القراءات بالربيع الأعلى ويرمز له بالرمز Q_3 ورتبته $\frac{3}{4}n$ كما يسمى المقدار الناتج من الفرق بين Q_3 و Q_1 بالمدى الربيعي وهو يمثل النصف الأوسط للقراءات ويؤخذ نصف هذا المدى مقياسًا للتشتت ويسمى بنصف المدى الربيعي ويرمز له بالرمز Q ويعطى بالعلاقة :

$$Q = \frac{Q_3 - Q_1}{2} \qquad \text{(٤ - ١)}$$

ويلاحظ أن Q_2 هو الربيع الثاني وهو القيمة التي يسبقها نصف البيانات ورتبته $\frac{n}{2}$ أي أن Q_2 هو الوسيط الذي سبق شرحه في مقاييس النزعة المركزية .

وسوف نتناول شرح نصف المدى الربيعي Q في كل من البيانات المباشرة والبيانات المبوبة كالتالي .

(٤ - ٣ - ١) نصف المدى الربيعي للبيانات المباشرة

يحسب كالتالي :

١) نرتب البيانات ترتيبًا تصاعديًا .

٢) نوجد قيمة Q_1 وهي القيمة التي يسبقها ربع القراءات .

٣) نوجد قيمة Q_3 وهي القيمة التي يسبقها ثلاثة أرباع القراءات .

وبتطبيق القانون : $Q = \frac{Q_3 - Q_1}{2}$

يتم حساب نصف المدى الربيعي .

مثال (٤ ـ ٣)

أوجد نصف المدى الربيعي لأوزان مجموعة الطلاب التالية :

67, 65, 69, 58, 55, 71, 72, 70

الحـــل

ترتب البيانات ترتيبًا تصاعديًّا نحصل على :

$$55\ ,\ 58 \uparrow 65\ ,\ 67,\ 69\ ,\ 70 \uparrow 71\ ,\ 72$$

$$Q_1 = \frac{58 + 65}{2} = 61.5\ ;\ Q_3 = \frac{70 + 71}{2} = 70.5$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{70.5 - 61.5}{2} = 4.5$$

مثال (٤ ـ ٤)

أوجد نصف المدى الربيعي لأوزان مجموعة الطلاب التالية :

59, 67, 65, 69, 58, 55, 70, 72, 74

الحـــل

بترتيب البيانات تصاعديًّا نحصل على :

$$50, 58, \boxed{59}, 65, 67, 69, \boxed{70}, 72, 74$$

$$Q_1 = 59 \quad , \quad Q_3 = 70$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{70 - 59}{2} = 5.5$$

(٤ ـ ٣ ـ ٢) نصف المدى الربيعي للبيانات المبوبة

يحسب نصف المدى لهذه البيانات بطريقتين الأولى حسابية والثانية بيانية وسوف نستعرض كل طريقة على حدة كمايلي .

نصف المدى الربيعي بالطريقة الحسابية

يتم حساب نصف المدى الربيعي بنفس الطريقة التي سبق شرحها لحساب الوسيط وهو طريقة الفروق «لبيرسون» ويحسب الربيع الأدنى Q_1 بوضع $\frac{n}{4}$ بدلاً من $\frac{n}{2}$ في قانون الوسيط ويحسب الربيع الأعلى بوضع $\frac{3n}{4}$ بدلاً من $\frac{n}{2}$ في قانون الوسيط وبعد ذلك نحسب نصف المدى الربيعي من العلاقة (٤ ـ ١) ويوضح طريقة حساب Q_1، Q_3 بالعلاقتين الآتيتين أي أن :

$$Q_1 = A_1 + \frac{\frac{n}{4} - f_1}{f_2 - f_1} \cdot L \quad \text{.........................} \quad (٤ - ٢)$$

$$Q_3 = A_2 + \frac{\frac{3n}{4} - f'_1}{f'_2 - f'_1} \cdot L \quad \text{.........................} \quad (٤ - ٣)$$

مثال (٤ ـ ٥)

أوجد نصف المدى الربيعي حسابيًّا لدرجات الطلّاب في مثال (٤ ـ ٢) .

الحــل

نكوّن الجدول المتجمع الصاعد كمايلي :

	التكرار المتجمع الصاعد	حـدود الفئات	
	0	أقــل مــن 39.5	
f_1	2	أقــل مــن 49.5	A_1
f_2	11	أقــل مــن 59.5	
f'_1	26	أقــل مــن 69.5	A_2
f'_2	37	أقــل مــن 79.5	
	39	أقــل مــن 89.5	
	40	أقــل مــن 99.5	

$$n = 40 \;,\; \frac{n}{4} = \frac{40}{4} = 10 \;,\; \frac{3n}{4} = \frac{120}{4} = 30 \;,\; L = 10$$

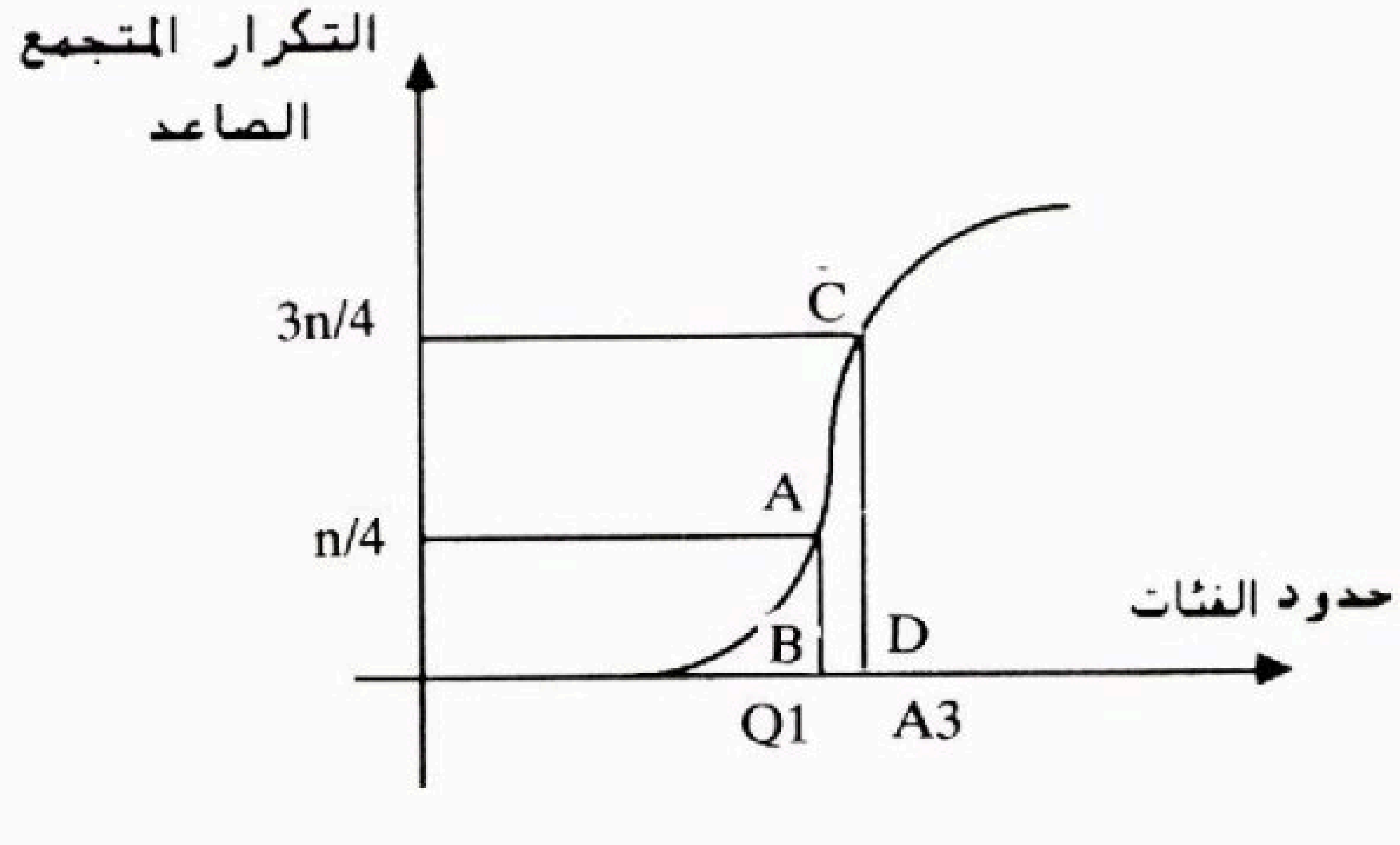
$$Q_1 = A_1 + \frac{\frac{n}{4} - f_1}{f_2 - f_1} . L = 49.5 + \frac{10-2}{11-2} \times 10 = 49.5 + 8.89 = 58.39$$

$$Q_3 = A_2 + \frac{\frac{3n}{4} - f_1'}{f_2' - f_1'} . L = 69.5 + \frac{30-26}{37-26} \times 10 = 69.5 + 3.64 = 73.14$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{73.14 - 58.39}{2} = 7.38$$

نصف المدى الربيعي بالطريقة البيانية

يُرسم المنحنى المتجمع الصاعد ثم يحدد على محور التكرارات المتجمعة الصاعدة $\frac{n}{4}$ ومنها يرسم خط يوازي محور الفئات حتى يقطع المنحنى المتجمع في نقطة A ومن A نسقط عمودًا رأسيًا يقابل محور الفئات في نقطة ولتكن B تكون القيمة عندها هي Q_1 ونحدد قيمة $\frac{3n}{4}$ على محور التكرارات المتجمعة الصاعدة ثم نرسم خطًا أفقيًا يقطع المنحنى المتجمع الصاعد في نقطة C ومن C نسقط عمودًا رأسيًا يقطع محور الفئات في نقطة D يكون عندها قيمة Q_3 ونحسب نصف المدى الربيعي باستخدام العلاقة (٤ ـ ١) ويوضح ذلك بالرسم، شكل (٤ ـ ١) التالي :



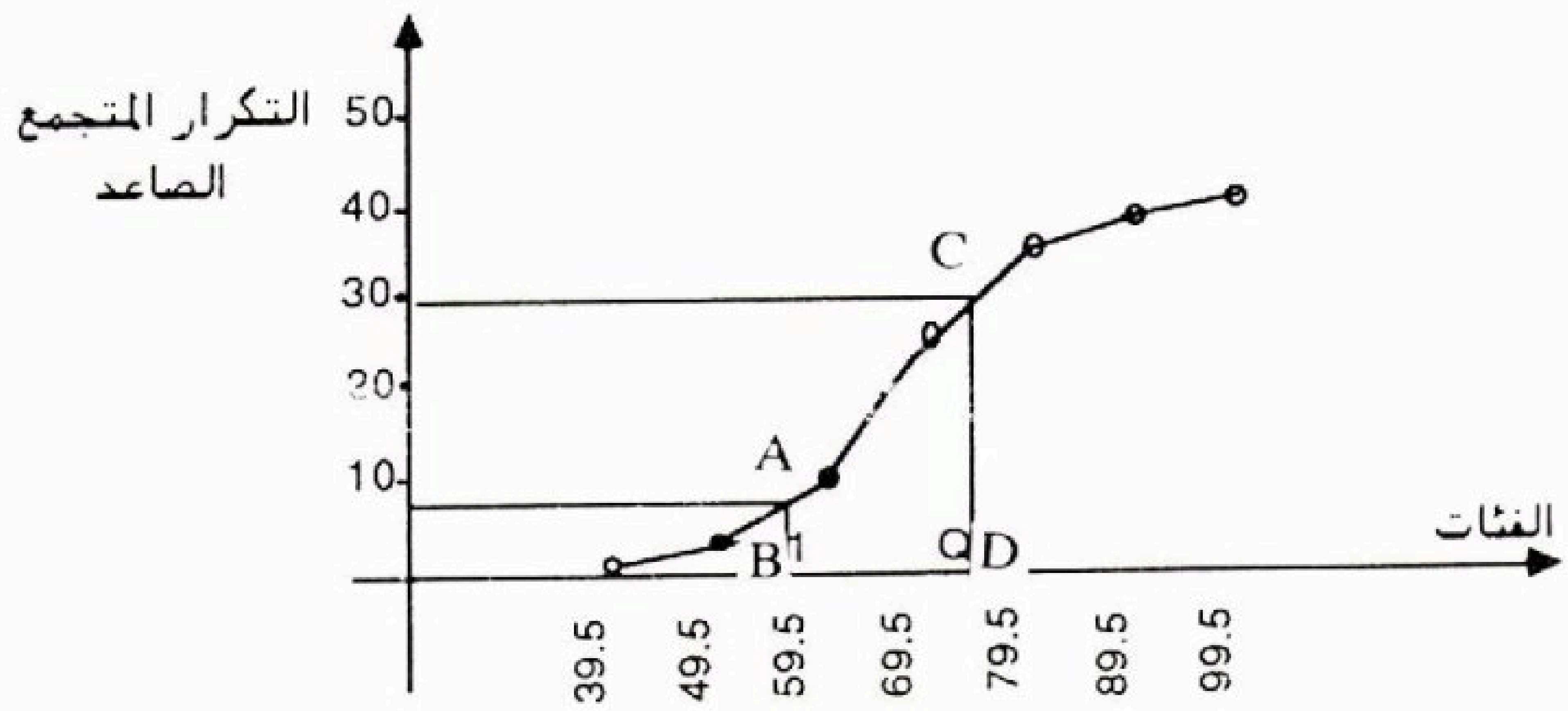
شكل (٤ ـ ١) يوضح قيمة كل من Q_1, Q_2 بيانيًا

مثال (٤ ـ ٦)

أوجد نصف المدى الربيعي لدرجات الطلّاب في مثال (٤ ـ ٢) السابق.

الحـــل

باستخدام الجدول المتجمع الصاعد في مثال (٤ ـ ٥) نرسم منه المنحنى المتجمع الصاعد ومن الرسم نوجد قيمة Q_3 و Q_1 ونحسب من العلاقة (٤ ـ ١) قيمة نصف المدى الربيعي Q كما في شكل (٤-٢) التالي: التكرار المتجمع



شكل (٤ ـ ٢) يبين قيمة Q_1, Q_3 لدرجات الطلاب بيانيًا

تقريبًا من الرسم: $Q_1 = 58$, $Q_3 = 70$

وعليه فإن

$$Q = \frac{70 - 58}{2} = \frac{12}{2} = 6$$ درجات

(٤ ـ ٣ ـ ٣) مميزات نصف المدى الربيعي

١) يتخلص من القيم المتطرفة الشاذة نحو الكبر أو الصغر.

٢) يمكن حسابه من التوزيعات التكرارية المفتوحة من طرف واحد أو من الطرفين.

(٤ ـ ٣ ـ ٤) عيوب نصف المدى الربيعي

١) لا يأخذ جميع القيم في الاعتبار.

٢) لا يسهل التعامل معه في التحليل الإحصائي.

(٤ ـ ٤) الانحـراف المتوسـط (Mean deviation)

يعرف الانحراف المتوسط بأنه متوسط الانحرافات المطلقة للقراءات عن وسطها الحسابي $\bar{x}$ ويرمز له بالرمز M.D ويعرّف رياضيًّا كالتالي :

$$M.D. = \frac{1}{n} \sum | x - \bar{x} | \qquad (٤ - ٤)$$

والسبب في أخذ القيم المطلقة للانحرافات هو أن مجموع انحرافات القيم عن وسطها الحسابي يساوي صفرًا (برهن ذلك). والتعريف السابق (٤ ـ ٤) يكون بالنسبة للبيانات المباشرة ولكن في حالة البيانات المبوبة يعطى الانحراف المتوسط بالعلاقة الآتية :

$$M.D = \frac{\sum f \, | x - \bar{x} |}{n} \qquad (٤ - ٥)$$

ونوضح ذلك بالأمثلة التالية .

مثال (٤ ـ ٧)

أوجد الانحراف المتوسط لأعمار مجموعة الطلاب

6, 5, 7, 7, 8, 9, 9, 5

الحــل

نكوّن جدول الحل كالتالي :

x	$x - \bar{x}$	$\lvert x - \bar{x} \rvert$
6	− 1	1
5	− 2	2
7	0	0
7	0	0
8	1	1
9	2	2
9	2	2
5	− 2	2
56	0	[illegible]

$$\bar{x} = \frac{\sum x}{n} = \frac{56}{8} = 7$$

وباستخدام العلاقة (٤ ـ ٤) نحصل على:

$$M.D. = \frac{10}{8} = 1.25$$ سنة

مثال (٤ ـ ٨)

أوجد الانحراف المتوسط لدرجات الطلّاب في مثال (٤ ـ ٢)

الحــل

يكون جدول الحل كما يلي:

	الفئـات	x	f	fx	$x-\bar{x}$	$\lvert x-\bar{x} \rvert$	$f\lvert x-\bar{x} \rvert$
	40 – 49	44.5	2	89	– 21.25	21.25	42.5
	50 – 59	54.5	9	490.5	– 11.25	11.25	101.25
	60 – 69	64.5	15	967.5	– 1.25	1.25	18.75
	70 – 79	74.5	11	819.5	8.75	8.75	96.25
	80 – 89	84.5	2	169.0	18.75	18.75	37.5
	90 – 99	94.5	1	94.5	28.75	28.75	28.75
Σ			40	2630			325

$$\bar{x} = \frac{\sum fx}{n} = \frac{2630}{40} = 65.75$$ درجة

$$M.D. = \frac{\sum f\lvert x-\bar{x} \rvert}{n} = \frac{325}{40} = 8.125$$ درجة

ملاحظـة

أحيانًا يعرّف الانحراف المتوسط باستخدام الوسيط بدلًا من الوسط الحسابي أو أي متوسطات أخرى غير الوسط الحسابي.

(٤ ـ ٥) الانحراف المعياري (Standard deviation)

من الصعب التعامل رياضيًّا (تحليليًّا) مع الانحراف المتوسط ولذلك دعت الحاجة إلى استخدام مقياس للتشتت بقوة الانحراف المتوسط نفسها أو أكثر ولكي يكون من السهل التعامل معه تحليليًّا وبما أن الفكرة هي التخلص من الإشارات للانحرافات فإن تربيع الانحراف يخلصنا من الإشارة ولهذا فإن الانحراف المعياري يعرّف عن طريق التباين الذي يعرَّف بأنه متوسط مربع انحرافات القيم عن وسطها الحسابي ويرمز له بالرمز σ^2 ويقرأ (سجما تربيع) والجذر التربيعي للتباين ينتج عنه مقياس من أهم وأدق مقاييس التشتت وهو ما يسمى بالانحراف المعياري ويرمز له بالرمز σ وسوف نتناول طرق حسابه في حالة البيانات المباشرة والبيانات المبوبة كما يلي.

(٤ ـ ٥ ـ ١) الانحراف المعياري للبيانات المباشرة

إذا كان لدينا قراءات من مجتمع إحصائي عدد مفرداته N هي

$$x_1, x_2, \ldots, x_N$$

ومتوسط هذه القراءات $(\bar{x})$ فإن مربعات انحرافات هذه القيم من $\bar{x}$ هي

$$(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \ldots, (x_N - \bar{x})^2$$

ويعرف التباين σ^2 كالتالي

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_N - \bar{x})^2}{N}$$

أي

$$\sigma^2 = \frac{1}{N} \sum (x - \bar{x})^2 \qquad \ldots\ldots\ldots\ldots \quad (٤ - ٦)$$

والانحراف المعياري هو

$$\sigma = \sqrt{\frac{1}{N} \sum (x - \bar{x})^2} \qquad \ldots\ldots\ldots\ldots \quad (٤ - ٧)$$

ويفضل عند حساب الانحراف المعياري σ أن يحسب التباين σ^2 من المعادلة (٤ ـ ٦) وبأخذ الجذر التربيعي للنتيجة النهائية نحصل على σ وفي حالة العينة التي

حجمها n المأخوذة من المجتمع فإن الانحراف المعياري في هذه الحالة يرمز له بالرمز s والتباين s^2 ويعرف بقسمة مجموع مربعات الانحراف على $(n - 1)$ بدلاً من (N) في حالة σ^2 ويكتب كما يلي:

$$s^2 = \frac{1}{n-1} \sum (x - \bar{x})^2 \quad \text{............} \quad (٤ - ٨)$$

وعليه فإن الانحراف المعياري s يكون

$$s = \sqrt{\frac{1}{n-1} \sum (x - \bar{x})^2} \quad \text{............} \quad (٤ - ٩)$$

وذلك في حالة البيانات المباشرة فإن s يعطي تقديرًا أحسن للانحراف المعياري σ للمجتمع الذي أخذت منه العينة. وإذا كان عدد مفردات العينة كبيرًا (أكبر من 30) فإن قيمة s^2، σ^2 متساويتان تقريبًا من الناحية العملية.

مثال (٤ - ٩)

أحسب الانحراف المعياري s لأعمار مجموعة من الطلاب في المرحلة الابتدائية وهي:

سنة 8, 9, 7, 6, 5

الحـــل

	x	$x-\bar{x}$	$(x-\bar{x})^2$
	8	1	1
	9	2	4
	7	0	0
	6	− 1	1
	5	− 2	4
Σ	35		10

حيث

$$\bar{x} = \frac{\sum x}{n} = \frac{35}{5} = 7$$ سنوات

$$s^2 = \frac{1}{n-1}\sum (x-\bar{x})^2 = \frac{10}{5-1} = \frac{10}{4} = 2.5$$

$$s = \sqrt{2.5 = 1.591}$$ سنة

ويمكن تبسيط العلاقة (٤ ـ ٨) السابقة كالتالي:

$$s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{(\sum x)^2}{n}\right] \quad \ldots\ldots\ldots\ldots \quad (١٠ ـ ٤)$$

الإثبـــات

$$s^2 = \frac{1}{n-1}\sum (x-\bar{x})^2 = \frac{1}{n-1}\sum (x^2 - 2x\bar{x} + \bar{x}^2)$$

$$= \frac{1}{n-1}(\sum x^2 - 2\bar{x}\sum x + n\bar{x}^2) = \frac{1}{n-1}(\sum x^2 - n\bar{x}^2)$$

$$= \frac{1}{n-1}(\sum x^2 - \frac{(\sum x)^2}{n})$$

ويلاحظ أن العلاقة (٤ ـ ١٠) تحتاج في الحساب إلى $\sum x^2$ ، $\sum x$ فقط.

مثال (٤ ـ ١٠)

أحسب الانحراف المعياري لأعمار الطلاّب في مثال (٤ ـ ٩) باستخدام العلاقة (٤ ـ ١٠).

الحـــل

نكوّن جدول الحل كما يلي :

x	x^2
8	64
9	81
7	49
6	36
5	25
35	255

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) = \frac{1}{4}\left(255 - \frac{(35)^2}{5}\right)$$

$$= \frac{1}{4}(255-245) = \frac{10}{4} = 2.5$$

$$s = \sqrt{2.5} = 1.581$$

وهي النتيجة نفسها في مثال (٤ ـ ٩) السابق.

(٤ ـ ٥ ـ ٢) بعض خصائص الانحراف المعياري

الخاصية الأولى

إذا أضفنا أو طرحنا مقدارًا ثابتًا c من جميع القراءات لمجموعة البيانات فإن الانحراف المعياري للقيم الجديدة هو الانحراف المعياري للقيم الأصلية نفسه ويمكن إثبات ذلك بالآتي :

نفرض أن القيم الأصلية هي :

$$x_1, x_2, \ldots, x_n$$

ونفرض أن القيم الجديدة هي :

$$d_1, d_2, \dots , d_n$$

حيث

$$d_1 = x_1 \pm c , d_2 = x_2 \pm c , \dots , d_n = x_n \pm c$$

أي أن

$$d = x \pm c$$

وحيث إن

$$s^2 = \frac{1}{n-1} \sum (x - \bar{x})^2$$

إذن

$$s^2 = \frac{1}{n-1} \sum [(d \pm c) - (\bar{d} \pm c)]^2 = \frac{1}{n-1} \sum (d - \bar{d})^2$$

أي أن

$$s^2 = \frac{1}{n-1} [\sum d^2 - \frac{1}{n} (\sum d)^2] \quad \dots\dots\dots \quad (٤ - ١١)$$

ويمكن أن تستخـدم هذه الخاصية في تبسيط القراءات وخاصة عندما تكون قيمتها كبيرة كما يتضح ذلك من المثال التالي .

مثال (٤ ـ ١١)

استخدم الخاصية الأولى السابقة في حل مثال (٤ ـ ٩) باختيار الثابت c يساوي 5 .

الحـــل

نطرح المقدار الثابت c = 5 من كل القراءات كما هو موضح بالجدول الآتي :

	x	d = x − 5	d^2
	8	3	9
	9	4	16
	7	2	4
	6	1	1
	5	0	0
Σ		10	30

وبالتعويض في (٤ - ١١) نحصل على:

$$s^2 = \frac{1}{n-1}\left(\sum d^2 - \frac{(\sum d)^2}{n}\right) = \frac{1}{4}\,(30 - \frac{(10)^2}{5}) = \frac{10}{4}$$

$$s^2 = 2.5 \quad , \quad s = \sqrt{2.5} = 1.581$$

وهي النتيجة نفسها التي في (٤ - ٩) السابق.

الخاصية الثانية

إذا ضربنا جميع القيم في مقدار ثابت أو قسمناها على مقدار ثابت فإن الانحراف المعياري يتأثر بذلك وسوف نثبت ذلك في حالة الضرب ويمكن اتباع الخطوات نفسها في حالة القسمة كما يلي:

نفرض أن القراءات هي $x_1, x_2, \ldots , x_n$

فإذا ضربنا هذه القيم في مقدار ثابت c تكون القيم الجديدة

$$d_1, d_2, \ldots , d_n$$

حيث

$$d_1 = cx_1 \, , \, d_2 = cx_2 \, , \, \ldots \, , \, d_n = cx_n$$

وعليه فإن

$$s_x^2 = \frac{1}{n-1} \sum (x - \bar{x})^2$$

$$s_x^2 = \frac{1}{n-1} \sum (\frac{d}{c} - \frac{d}{c})^2 = \frac{1}{c^2}\,\frac{1}{n-1} \sum (d - \bar{d})^2$$

$$s_x^2 = \frac{1}{c^2}\, s_d^2 \quad , \quad s_x = \frac{1}{c}\, s_d \qquad \text{(٤ - ١٢)}$$

أي أن الانحراف المعياري للقيم الأصلية في حالة الضرب يساوي الانحراف المعياري للقيم الجديدة مقسومًا على المقدار الثابت كما هو موضح بالعلاقة (٤ - ١٢) أما في حالة القسمة فإنه يمكن إثبات أن

$$s_x = |\,c\,|\, s_d \qquad \text{(٤ - ١٣)}$$

حيث $|c|$ القيمة المطلقة للثابت c (الموجبة).

أي أن الانحراف المعياري للقيم الأصلية يساوي الانحراف المعياري للقيم الجديدة مضروبًا في المقدار c كما هو موضح بالعلاقة (٤ - ١٣).

الخاصية الثالثة

مجموع مربعات الانحراف للقيم عند وسطها الحسابي $\bar{x}$ تكون أصغر من مجموع مربعات الانحراف للقيم عن أي وسط فرضي آخر a حيث $a \neq \bar{x}$.

الإثبات

$$\begin{aligned}
\sum (x-a)^2 &= \sum (x+\bar{x}-\bar{x}-a)^2 \\
&= \sum [\,(x-\bar{x}) + (\bar{x}-a)\,]^2 \\
&= \sum (x-\bar{x})^2 + n\,(\bar{x}-a)^2 + 2\,(\bar{x}-a) \sum (x-\bar{x}) \\
&= \sum (x-\bar{x})^2 + n\,(\bar{x}-a)^2
\end{aligned}$$

ونلاحظ أن المقدار $n\,(\bar{x}-a)^2$ مقدار موجب دائمًا ونستنتج من ذلك أن

$$\sum (x-\bar{x})^2 < \sum (x-a)^2$$

وهو المطلوب.

الخاصية الرابعة

إذا كانت هناك عينتان مجموع تكرارهما هو n_1, n_2 وتباينهما هو s_1^2, s_2^2 على الترتيب ولهما المتوسط $\bar{x}$ نفسه فإن التباين المشترك هو:

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-1}$$

الإثبات

نفرض المجموعتين هما $x_1, \dots, x_{n_1}\ ;\ y_1, \dots, y_{n_2}$

$$s_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i-\bar{x})^2 \ , \ s_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_i-\bar{x})^2$$

$$s_1^2(n_1-1) = \sum_{i=1}^{n_1}(x_i-\bar{x})^2$$

$$s_2^2(n_2-1) = \sum_{i=1}^{n_2}(y_i-\bar{x})^2$$

$$s_1^2(n_1-1)+s_2^2(n_2-1) = \sum_{i=1}^{n_1}(x_i-\bar{x})^2 + \sum_{i=1}^{n_2}(y_i-\bar{x})^2$$

$$= \sum_{i=1}^{n_1+n_2}(x_i-\bar{x})^2 \ ; \ x_i = y_i \quad , \quad i = n_1+1, \dots , n_1+n_2$$

$$s^2 = \frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-1}$$

الخاصية الخامسة

الانحراف المعياري لمجموعة من البيانات أكبر من الانحراف المتوسط لها (على الطالب التحقق من ذلك).

(٤ ـ ٥ ـ ٣) الانحراف المعياري للبيانات المبوبة

في هذه الحالة إذا كانت لدينا تكرارات $f_1, f_2, \dots , f_k$ ولها مراكز فئات $x_1, x_2, \dots ,$ x_k على الترتيب فإن المعادلات السابقة (٤ ـ ٨)، (٤ ـ ٩)، (٤ ـ ١٠)، (٤ ـ ١١) السابقة تصبح كالتالي:

$$s^2 = \frac{1}{n-1}\sum f(x-\bar{x})^2 \ , \ s = \sqrt{\frac{1}{n-1}\sum f(x-\bar{x})^2} \quad \dots\dots \quad (٤ ـ ١٤)$$

$$s^2 = \frac{1}{n-1}\left(\sum fx^2 - \frac{(\sum fx)^2}{n}\right) \quad \dots\dots \quad (٤ ـ ١٥)$$

$$s^2 = \frac{1}{n-1}\left(\sum fd^2 - \frac{(\sum fd)^2}{n}\right) \quad \dots\dots \quad (٤ ـ ١٦)$$

وسـوف نبين طريقة حساب الانحراف المعياري باستخدام العلاقات السابقة (٤ ـ ١٤)، (٤ ـ ١٥)، (٤ ـ ١٦) بالأمثلة التالية .

مثال (٤ ـ ١٣)

أوجد الانحراف المعياري لدرجات الطلاب في مثال (٤ ـ ٢) وذلك باستخدام العلاقات (٤ ـ ١٤)، (٤ ـ ١٥)، (٤ ـ ١٦)

الحـــل

(i) باستخدام العلاقة (٤ ـ ١٤) يكون جدول الحل كالتالي :

الفئــات	x	f	fx	$x-\bar{x}$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
40–49	44.5	2	89.0	− 21.25	451.56	903.13
50–59	54.5	9	490.5	− 11.25	126.56	1139.06
60–69	64.5	15	967.5	− 1.25	1.56	23.44
70–79	74.5	11	819.5	8.75	76.56	842.19
80–89	84.5	2	169.0	18.75	351.56	703.13
90–99	94.5	1	94.5	28.75	826.56	826.56
		40	2630			4437.5

$$\bar{x} = \frac{1}{n} \sum fx = \frac{1}{40}(2630) = 65.75$$

$$s^2 = \frac{1}{n-1} \sum f(x-\bar{x})^2 = \frac{1}{40-1}(4437.51) = 113.78$$

$$s = 10.67$$ درجة

(ii) باستخدام العلاقة (٤ ـ ١٥) يكون جدول الحل كالتالي :

الفئـات	x	f	fx	fx^2
40 – 49	44.5	2	89.0	3960.50
50 – 59	54.5	9	490.5	26732.25
60 – 69	64.5	15	967.5	62403.75
70 – 79	74.5	11	819.5	61052.75
80 – 89	84.5	2	169.0	12280.5
90 – 99	94.5	1	94.0	8930.25
			2630	177360

$$s^2 = \frac{1}{n-1}\left(\sum fx^2 - \frac{(\sum fx)^2}{n}\right) = \frac{1}{39}(177360 - 172922.5)$$

$$s^2 = 113.78 \quad , \quad s = \sqrt{113.78} = 10.67$$ درجة

وهي النتيجة السابقة نفسها .

(iii) باستخدام العلاقة (٤ ـ ١٦) وبأخذ المقدار الثابت (الوسط الفرضي) $c = 64.5$ وهو مركز الفئة التي يقابلها أكبر تكرارًا (وذلك لتبسيط الحسابات) كما هو موضح بجدول الحل التالي :

الفئـات	x	f	d = x – 64.5	fd	fd^2
40 – 49	44.5	2	– 20	– 40	800
50 – 59	54.5	9	– 10	– 90	900
60 – 69	64.5	15	0	0	0
70 – 79	74.5	11	10	110	1100
80 – 89	84.5	2	20	40	800
90 – 99	94.5	1	30	30	900
		40		50	4500

$$s^2 = \frac{1}{n-1}\left(\sum fx^2 - \frac{(\sum fx)_2}{n}\right) = \frac{1}{39}(4500 - 62.5)$$

$$s^2 = 113.78 \quad , \quad s = \sqrt{113.78} = 10.67$$ درجة

ملاحظة

باستخـدام الخاصية الثانية في حل المثال السابق، لقد لاحظنا أنه باستخدام الـوسط الفـرضي وهو مركز الفئة التي يقابلها أكبر تكرار قد بسطت الحسابات كثيراً ويمكن تبسيط الحسابات أكثر وذلك بقسم انحرافات القيم عن الوسط الفرضي على طول الفئة (وتستخدم هذه الطريقة في حالة الفئات المنتظمة) وبذلك يصبح الحل كما يلي:

الفئـــات	x	f	$d = x - 64.5$	$d' = \frac{d}{10}$	fd'	fd'^2
40 – 49	44.5	2	– 20	– 2	– 4	8
50 – 59	54.5	9	– 10	– 1	– 9	9
60 – 69	64.5	15	0	0	0	0
70 – 79	74.5	11	10	1	11	11
80 – 89	84.5	2	20	2	4	8
90 – 99	94.5	1	30	3	3	9
		40			5	45

$$s^2_{d'} = \frac{1}{n-1}\left(\sum fd'^2 - \frac{(\sum fd')^2}{n}\right)$$

$$= \frac{1}{39}\left(45 - \frac{(5)^2}{40}\right) = 1.1378$$

$$s_{d'} = 1.067 \quad , \quad s_x = 10 s_{d'} = 10.67$$ درجة

وهي النتيجة السابقة نفسها.

مميزات الانحراف المعياري وعيوبه هي مميزات الوسط الحسابي وعيوبه نفسها التي سبق ذكرها في الفصل الثالث.

(٤ ـ ٦) معامل الاختلاف (Coefficient of variation)

من المعلوم أن الانحراف المعياري لمجموعة من البيانات يأخذ وحدات البيانات نفسها فإذا كانت البيانات تمثل الأطوال مقيسة بالسنتيمترات فإن الانحراف المعياري يكون بالسنتيمتر وإذا كانت البيانات تمثل الأوزان فإنها تكون مقيسة بالكيلوجرام ويكون الانحراف المعياري مقيسًا بالكيلوجرام فإذا أردنا مقارنة تجانس مجموعة من الأوزان أو تشتتها بمجموعة من الأطوال فلا يمكن استخدام الانحراف العياري للمقارنة لأنه لا يمكن مقارنة السنتيمتر بالكيلوجرام لذا دعت الحاجة إلى إيجاد مقياس لا يعتمد على الوحدات وهذا المقياس هو ما يسمى بمعامل الاختلاف أو التشتت النسبي ويعرف كالتالي:

$$C.V = \frac{s}{\bar{x}} \qquad \text{.........................} \qquad (٤ ـ ١٧)$$

$$C.V = \frac{Q_3 - Q_1}{Q_3 + Q_1} \qquad \text{.........................} \qquad (٤ ـ ١٨)$$

مثال (٤ ـ ١٤)

أحسب معامل الاختلاف لدرجات الطلاب في مثال (٤ ـ ١٣) السابق بطريقتين مختلفتين.

الحـــل

الطريقة الأولى

سبق حساب المتوسط $\bar{x}$ والانحراف المعياري s في مثال (٤ ـ ١٣) حيث

$s = 10.67 \quad , \quad \bar{x} = 65.75$ درجة

باستخدام العلاقة (٤ ـ ١٧) السابقة نحصل على:

$$C.V = \frac{s}{\bar{x}} = \frac{10.67}{65.75} = 0.167$$

الطريقة الثانية

سبق حساب Q_1 (الربيع الأدنى)، Q_3 الربيع الأعلى في مثال (٤ ـ ٥) السابق حيث

$$Q_1 = 58.39 \quad , \quad Q_3 = 73.14 \quad \text{.......................} \quad (٤ - ٥)$$

وباستخدام العلاقة (٤ ـ ١٨) السابقة نحصل على:

$$C.V = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{73.14 - 58.39}{73.14 + 58.39} = \frac{14.75}{132.53} = 0.111$$

المقياس C.V يعطي درجة التفاوت بين المفردات ولا يعتمد على الوحدات.

(٤ ـ ٧) نظرية تشيبشيف

إذا كان لدينا مجموعة من البيانات ذات متوسط $\bar{x}$ وانحراف معياري s فإنه توجد نسبة تساوي على الأقل $1 - \frac{1}{k^2}$ من هذه البيانات (حيث k قيمة عددية أكبر من الواحد) تقع بين $(\bar{x} - ks \ ، \ \bar{x} + ks)$.

فمثلاً إذا كان لدينا مجموعة من البيانات لها متوسط 7 وانحراف معياري 5 فإنه لقيمة $k = 2$ يمكن القول إن 75% على الأقل من هذه البيانات تقع بين $(-3, 17)$. وهذه النظرية تستخدم لوصف البيانات التي ليس معلومًا عنها سوى المتوسط والانحراف المعياري (أنظر التمارين).

(٤ ـ ٨) المتغير المعياري والدرجات المعيارية (مقياس التمركز)

إذا كان لدينا المتغير X له القيم $x_1, x_2, \ldots, x_n$ والتي متوسطها $\bar{x}$ وانحرافها المعياري s فإن المتغير z والذي يعطى بالعلاقة التالية:

$$z_i = \frac{x_i - \bar{x}}{s} \quad , i = 1, 2, \ldots , n \qquad \text{(٤ - ١٩)}$$

حيث z_i تقيس الانحـرافـات عن المتوسط الحسابي بوحدات من الانحراف المعياري يسمى بالمتغير المعياري ويستخدم للمقارنة بين التوزيعات المختلفة ويوضح ذلك بالمثال التالي.

مثال (٤ ـ ١٥)

حصل طالب على 82 درجة في مقرر للإحصاء حيث كان متوسط الدرجات هو 75 درجة وانحراف معياري 10 درجات ثم حصل على 89 درجة في مقرر للرياضيات وكـان متوسط الدرجات للرياضيات هو 81 درجة وانحراف معياري 16 درجة في أي المقررين كانت درجة استيعاب هذا الطالب أعلى؟

الحـــل

إذا كانت z_1 ترمز للدرجة المعيارية للإحصاء فإن :

$$z_1 = \frac{82 - 75}{10} = 0.7$$

وإذا كانت z_2 ترمز للدرجة المعيارية للرياضيات فإن :

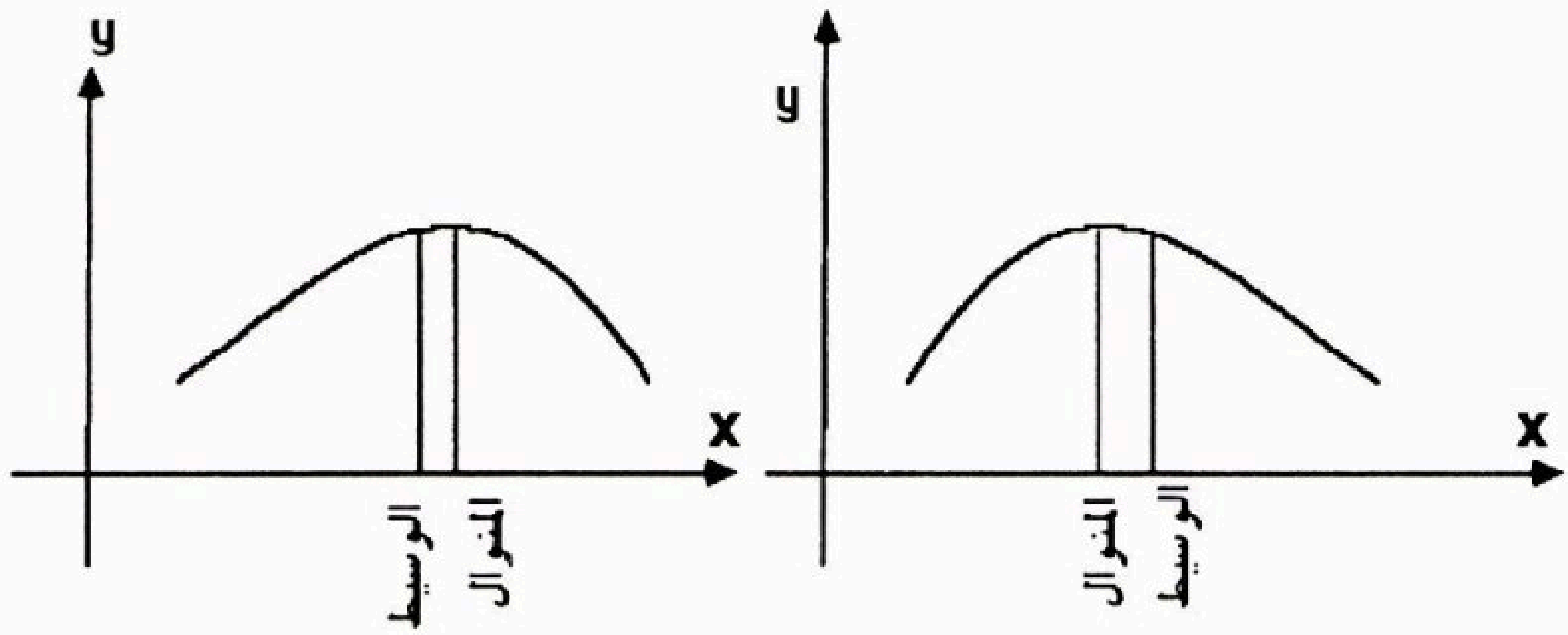
$$z_2 = \frac{89 - 81}{16} = 0.5$$

وهذا يعطي أن استيعاب الطالب النسبي لمقرر الإحصاء أعلى من الرياضيات.

(٤ ـ ٩) مقاييس الالتواء (الشكل) Skewness

لقد سبق أن تكلمنا عن طرق عرض البيانات جدوليًا وبيانيًا ثم مقاييس النزعة المـركـزية، ومقـاييس التشتت ولم تتعـرض لأي مقياس يوضح درجة التواء المنحنى التكـراري. ويقصـد بكلمـة الالتـواء هو بعد المنحنى عن التماثل، ويكون منحنى

التوزيع التكراري ملتويًا نحواليمين، فإن القيم المتطرفة نحو اليمين تؤثر على الوسط الحسابي وتسحبه نحو اليمين وبذلك يكون الوسط الحسابي أكبر من الوسيط، أما إذا كان التوزيع ملتويًا نحو اليسار فإن الوسط الحسابي يكون أصغر من الوسيط كما هو موضح بشكل (٤ ـ ٣) التالي:



شكل (٤ ـ ٣) يبين الالتواء نحو اليمين ونحو اليسار

مقياس الالتواء (v) له صور مختلفة نذكر ثلاثًا منها هي:

$$v = \frac{3(\bar{x} - med)}{s} \qquad \text{.........................} \qquad (٤ ـ ٢٠)$$

أو

$$v = \frac{(\bar{x} - mod)}{s} \qquad \text{.........................} \qquad (٤ ـ ٢١)$$

وتستخدم طريقة العزوم وتعطى بـ

$$v = \frac{m_3}{s^3} \qquad \text{.........................} \qquad (٤ ـ ٢٢)$$

حيث

بيانات مباشرة $$m_3 = \frac{\sum(x-\bar{x})^3}{n}$$

بيانات مبوبة $$m_3 = \frac{\sum f(x-\bar{x})^3}{n}$$

مثال (٤ ـ ١٦)

أوجد معامل الالتواء لدرجات الطلّاب في مثال (٤ ـ ١٣) .

الحـــل

سبق الحساب لدرجات الطلّاب القيم التالية :

$$\bar{x} = 65.75,\ \text{Mod} = 65.5,\ \text{Med} = 65.5,\ s = 10.67$$

معامل الالتواء من العلاقة (٤ ـ ٢٠) هو $v = \dfrac{3(65.75 - 65.5)}{10.67} = 0.070$

معامل الالتواء من العلاقة (٤ ـ ٢١) هو $v = \dfrac{65.75 - 65.5}{10.67} = 0.023$

(٤ ـ ١٠) التفلطح (Kurtosis)

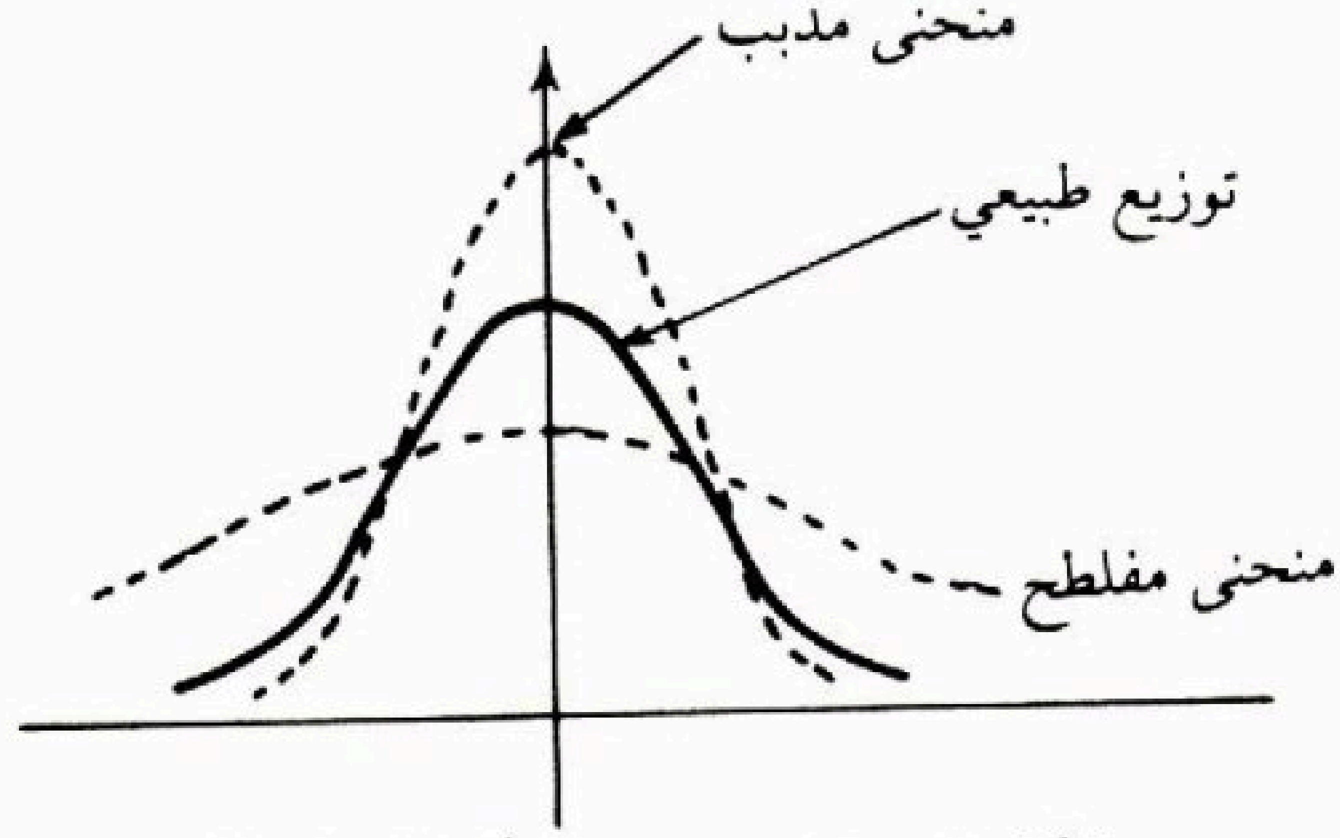
هو مقياس يقيس تدبب (أو الاستواء) لمنحنى الكثافة عند المتوسط، وهو أيضًا مقياس يقيس درجة علو أي منحنى توزيع تكراري أو انخفاضه بالنسبة للمنحنى الطبيعي وهو منحنى متماثل حول محور رأسي يمر بالمتوسط ويُعرّف التفلطح k كما في شكل (٤ ـ ٤) التالي :

$$k = \frac{m_4}{s^4} \quad \text{.......................} \quad (٤ - ٢٣)$$

حيث إن :

(بيانات مباشرة) $m_4 = \dfrac{\sum (x - \bar{x})^4}{n}$

(بيانات مبوبة) $m_4 = \dfrac{\sum f(x - \bar{x})^4}{n}$



شكل (٤ ـ ٤) يبين بعض أشكال التفلطح

(٤ ـ ١١) تمــارين

(١) أحسب المدى ونصف المدى الربيعي والانحراف المتوسط والانحراف المعياري ومعامل الاختلاف من البيانات التالية:

$$6, 3, 5, 5, 9, 4, 6, 7, 1, 2, 4, 8$$

(٢) فيما يلي توزيع أوزان 50 طالبًا من طلاب جامعة الملك سعود.

73 – 75	70 – 72	67 – 69	64 – 66	61 – 63	58 – 60	**فئات الوزن**
4	8	15	14	7	2	**عدد الطلاب**

أوجد:

ا) مدى أوزان الطلاب.

ب) نصف المدى الربيعي للأوزان حسابيًا وبيانيًا.

جـ) الانحراف المتوسط والانحراف المعياري.

(٣) حدد نسبة الطلبة في المسألة (٢) والتي تقع أوزانهم في المدى

(i) $\bar{x} \pm 1\,(s)$, (ii) $\bar{x} \pm 2\,(s)$, (iii) $\bar{x} \pm 3\,(s)$

(٤) إذا كانت d هي انحرافات مراكز فئات لمجموعة من البيانات عن مقدار ثابت c أثبت أن الانحراف المعياري هو:

$$s = \sqrt{\frac{1}{n-1}\left(\sum fd^2 - \frac{(\sum fd)^2}{n}\right)}$$

(٥) أحسب مقاييس الالتواء ومقاييس التفلطح من البيانات في المسألة (٢).

(٦) مصنـع ينتج نوعين من لمبات التليفزيون هما B و A ومتوسط أعمارها الإنتاجية بالساعة هو $\bar{x}_B = 1500$ و $\bar{x}_A = 1200$ والانحراف المعياري بالساعة هو $s_B = 300$ و $s_A = 250$ أي من النوعين أكثر تشتتًا.

(٧) الجدول التالي يمثل دخل مجموعة من الأسر بمئات الريالات:

فئات الدخل	أقل من 10	10 - 14	15 - 19	20 - 24	25 - 29	فأكثر 30
عدد الأسر	5	20	35	19	13	8

أي من المقاييس التالية يمكن إيجادها وأي منها لا يمكن إيجادها مع ذكر السبب؟
المدى ـ نصف المدى الربيعي ـ معامل الاختلاف

(٨) أوجد العزم الأول والثاني والثالث والرابع لمجموعة البيانات:

2, 5, 9, 4, 3, 6

ثم أحسب معدل الالتواء ومعامل التفلطح لهذه البيانات.

(٩) عند دراسة أطوال مجموعة من الأطفال حديثي الولادة كانت أطوالهم هي:

70, 70, 70, 70, 70, 70, 70

أحسب مقاييس التشتت لهذه الأطوال.

(١٠) أخذت عيّنتان من مجتمعين فأعطتا النتائج التالية:

العيّنة الثانية	العيّنة الأولى
$\sum_{i=1}^{40} Y_i = 280$	$\sum_{i=1}^{50} X_i = 300$
$\sum_{i=1}^{40} Y_i^2 = 2100$	$\sum_{i=1}^{50} X_i^2 = 1950$

ا) أوجد الوسط الحسابي والانحراف المعياري لكل من العيّنتين.
ب) أي من العيّنتين أكثر تجانسًا.
جـ) إذا دمجت العيّنتان ما هو الوسط الحسابي للمجموعة الناتجة.

(١١) عنـد دراسـة ظاهـرة الطول والوزن لمجموعة عمال بأحد المصانع كانت لدينا البيانات التالية:

(i) ظاهرة الطول: $\bar{x} = 160\ \text{cm} \quad , \ s = 8\ \text{cm}$

(ii) ظاهرة الوزن: $\sum_{i=1}^{20} x_i = 1200\ \text{kg} \ , \ \sum_{i=1}^{20} x_i^2 = 72687\ \text{kg}.$

أيا من الظاهرتين أكثر تجانسًا.

(١٢) وجـد أن متـوسط كمية فيتـامين C في نوع معين من الفواكه هو 0.24 ملجم بانحراف معياري قدره 0.004 ملجم. فما هي أقل نسبة من الفاكهة التي تحتوي على مقدار من هذا الفيتامين واقع بين (0.248 , 0.232 ملجم).

(١٣) في دراسة قام بها مركز للأغذية وجد أن متوسط كمية فيتامين B في شرائح الخبز هو 0.260 ملجرام بانحراف قدره 0.005 ملجرام. أوجد القيم التي تقع بينها كمية فيتامين B في:

ا ـ على الأقل نسبة $\frac{35}{36}$ من هذه الشرائح.

ب ـ على الأقل نسبة $\frac{63}{64}$ من هذه الشرائح.

(١٤) ادّعت شركة طيران أن سفرياتها بين المدن الداخلية تصل متأخرة عن موعدها بمتوسط قدره 4.6 دقيقة وانحراف معياري مقداره 1.6 دقيقة. فما هي أقل نسبة من سفرياتها تصل متأخرة ما بين (1.8 و 7.4 دقيقة).

(١٥) حول مجموعة القيم التالية :

6, 5, 7, 2, 3, 9

إلى درجات معيارية .

(١٦) اثبت أن متوسط مجموعة من الدرجات المعيارية هو صفر وانحرافها المعياري هو واحد وضح ذلك باستخدام المسألة رقم (١٥) السابقة .

(١٧) ضع الإشارة (√) أمام العبارة الصحيحة والإشارة (X) أمام العبارة الخاطئة :

ا - تباين مجموعة من القياسات السالبة هو تباين سالب ☐

ب - يتأثر المدى وبشدة بالقيم الشاذة ☐

جـ - إذا كان المتغير X يعبر عن أطوال مجموعة من الأشخاص بالسم فإن قياس معامل الاختلاف بالسم أيضا ☐

د - تباين لمجموعة القيم 9، 9، 9، 9، 9 هو 81 ☐

(١٨) أكمل ما يلي كي تحصل على عبارات صحيحة :

ا - إذا علمت بيان إحصائي لدرجات اختبار يساوي 165 درجة وانحرافه المعياري 5 درجات فيها لا يقل عن ثلاثة أرباع الطلبة نالوا درجات واقعة بين (............ و ....) .

ب - بعد معايرة القياسات 15، 8، 81، 19، 20 يصبح متوسط القيم المعيارية .... ويصبح انحرافها المعياري ....

جـ - إذا كان متوسط مجموعة القياسات x_1, x_2, x_3, x_4 وانحرافها المعياري هما على الترتيب $\bar{x} = 15$ و $s = 3$ فإن المتوسط الحسابي والانحراف المعياري ومعامل الاختلاف للمجموعة $4x_1 - 5, 4x_2 - 5, 4x_3 - 5, 4x_4 - 5$ هي على الترتيب: المتوسط = ... والانحراف المعياري = ... ومعامل الاختلاف = ...

د - الوسيط لمجموعة القياسات 9-، 9، 2، 3-، 7- هو .... ومداها هو ....

٥

الارتباط والانحدار
Correlation and Regression

● مقدمة ● معامل الارتباط الخطي لبيرسون ● معامل الارتباط للرتب لسبيرمان ● معامل الاقتران ومعامل التوافق ● خط الانحدار ● تمارين

(٥ - ١) مقدمة

تناولنا في الفصول السابقة طرق دراسة متغير واحد لأي ظاهرة محل دراسة، مثل الأوزان لمجموعة من الطلّاب أو الأجور لمجموعة من العمّال أو... الخ وعرضنا كيف يمكن تلخيص بياناتها في جداول توزيعات تكرارية وعرضها بيانيًا وذلك للتعرف على خصائص المنحنيات التكرارية لها. وكذلك درسنا بعض المقاييس العددية التي تساعد على معرفة بعض خصائص التوزيعات التكرارية، ومن هذه المقاييس المتوسطات، والتشتت، والالتواء، والتفلطح. والآن سوف نتناول دراسة البيانات التي يكون لأفرادها متغيران يتغيران معًا في وقت واحد، وذلك لمعرفة نوع العلاقة التي تربط بينهما. والأمثلة كثيرة على هذا النوع من البيانات، مثل دراسة العلاقة بين أوزان مجموعة من الطلّاب وأطوالها، أو أعمار مجموعة من الطلّاب ودرجاتها أو أجور مجموعة من العمّال وإنتاجها، أو الدخل والإنفاق لمجموعة من الأسر، أو العلاقة بين صفة الطول للأب والابن، أو صفة الذكاء للأب والابن وهكذا. ثم إيجاد مقاييس تقيس درجة هذه العلاقة.

وكذلك في هذا الفصل أيضًا نتناول دراسة العلاقة بين المتغيرين (X, Y) ، فإذا كان هناك علاقة بين المتغير X والمتغير Y ، فكيف يمكن التعبير عنها بمعادلة رياضية ومنها يمكن التنبؤ بقيمة أحد المتغيرين إذا علمت قيمة المتغير الآخر.

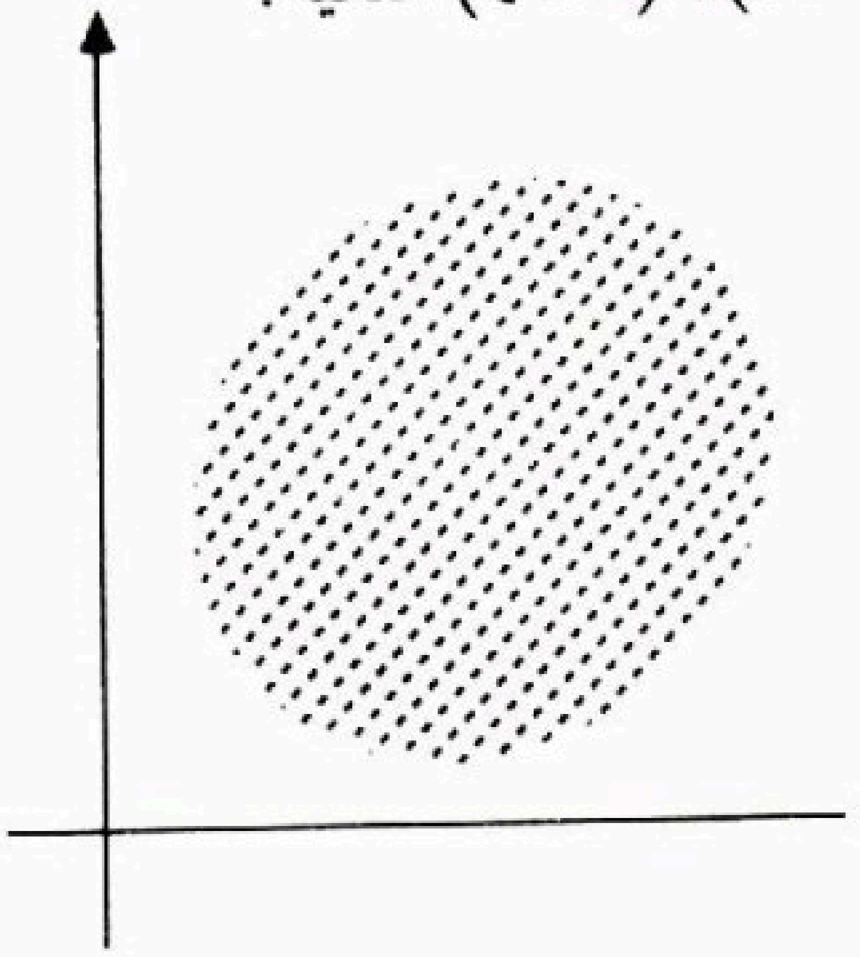
فإذا أردنـا دراسـة العلاقة بين الطول (Y) والوزن (X) لمجموعة عددها n من طلاب جامعة الملك سعود. فإنه يمكن لكل طالب قياس طوله ووزنه ويصبح لدينا أزواج من القراءات التالية :

$$(x_1, y_1), (x_2, y_2), \ldots\ldots , (x_n, y_n)$$

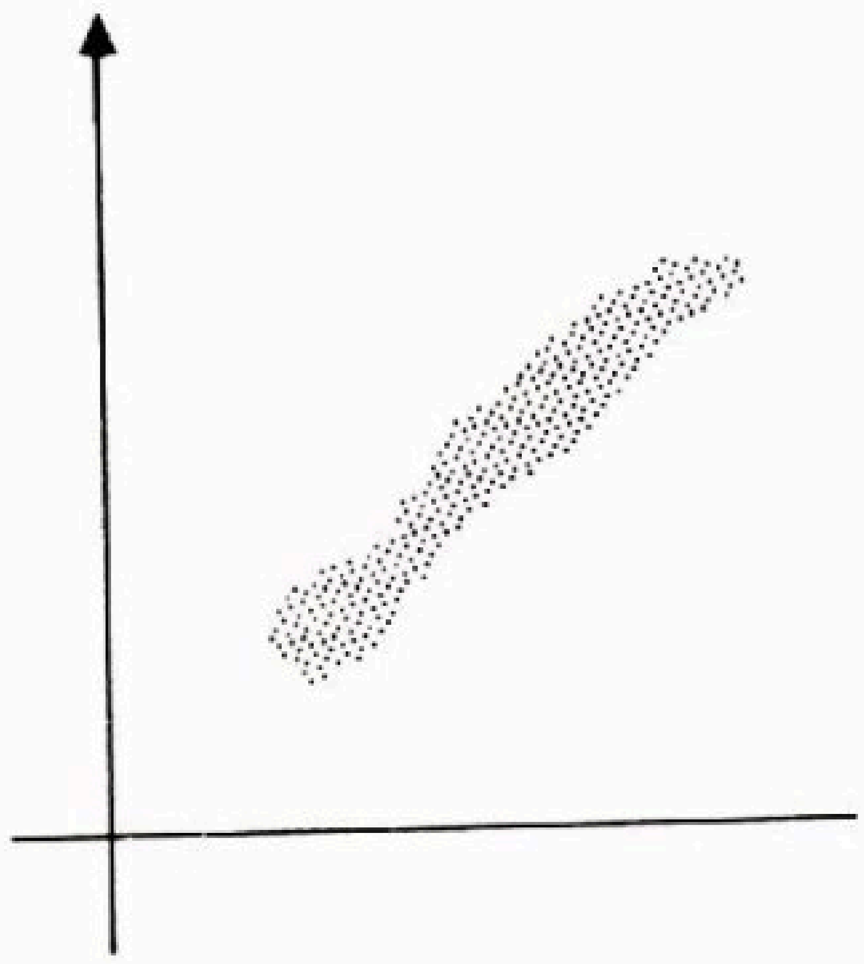
ولتمثيل هذه الأزواج من القـراءات بيانيًا نرسم محورين متعـامدين. المحور الأفقي ويمثل الوزن X مثلاً والمحور الرأسي ويمثل الطول Y ونقوم بتمثيل القراءات السابقة بنقاط فنحصل على ما يسمى بشكل الانتشار (scatter diagram) .

(٥ - ١ - ١) أشكال الانتشار

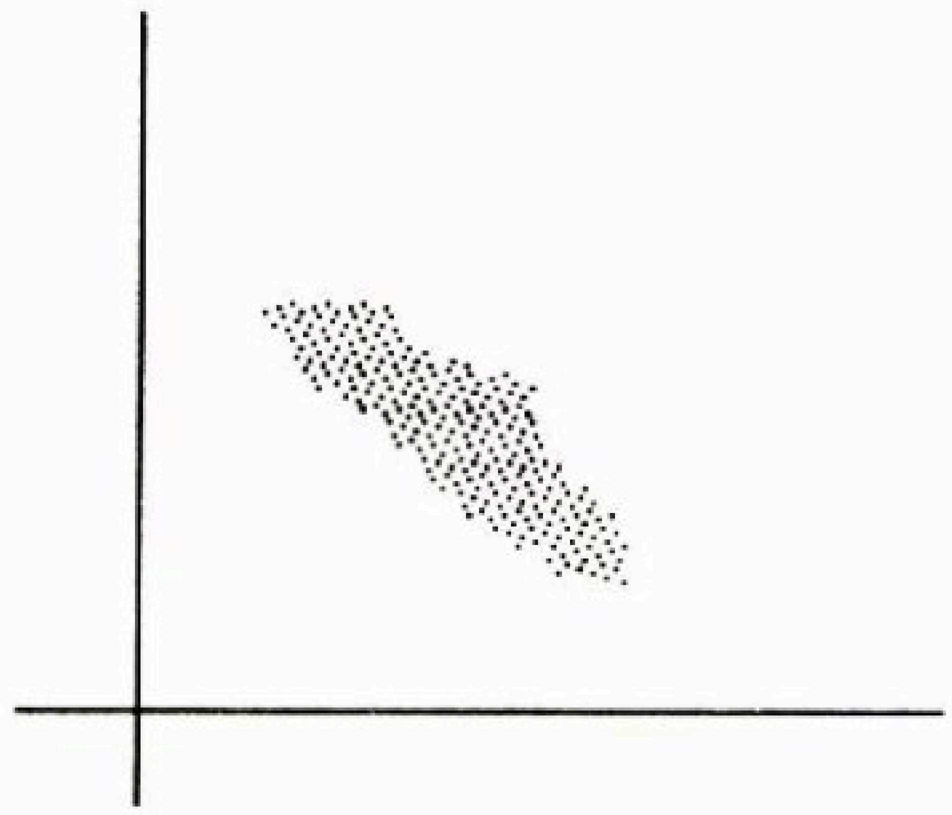
وأشكال الانتشار تأخذ صورًا مختلفة وذلك حسب طبيعة العلاقة بين المتغيرين (X, Y) محل الدراسة. وفيما يلي نعرض بعض أشكال الانتشار (٥ - ١)، (٥ - ٢)، (٥ - ٣)، (٥ - ٤) التالية :



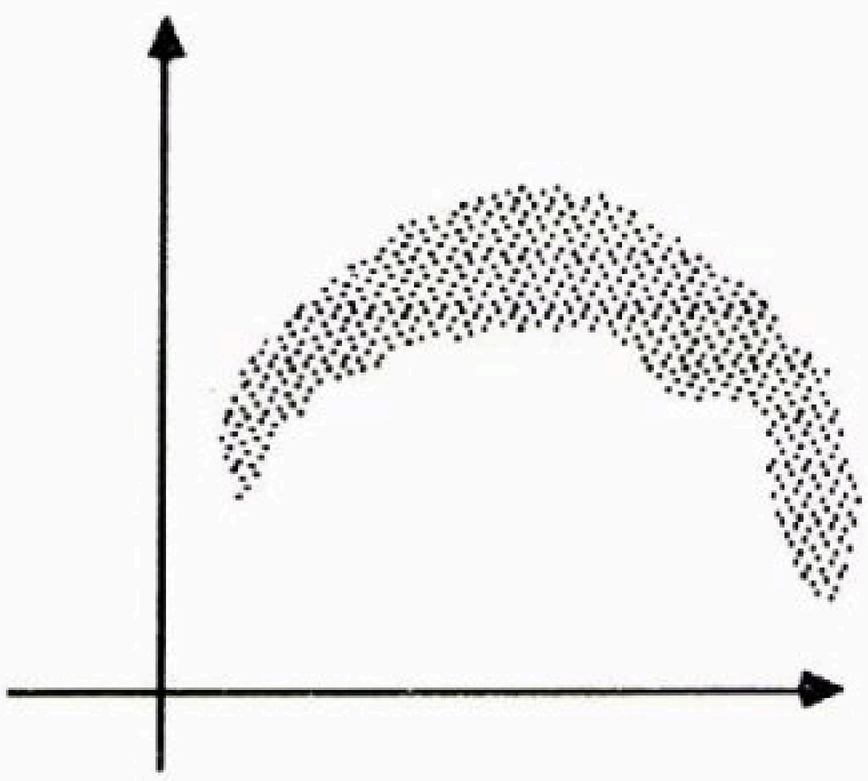
شكل (٥ - ١)



شكل (٥ - ٢)



شكل (٥ - ٣)



شكل (٥ - ٤)

ونلاحظ من أشكال الانتشار السابقة ما يلي :

شكل الانتشار (٥ - ١)

تكون فيه النقاط منتشرة بدون ترابط حول اتجاه محدد مما يدل على أنه لا توجد علاقة بين المتغيرين (X, Y) .

شكل الانتشار (٥ - ٢)

تكون فيه النقاط منتشرة حول خط مستقيم تزيد فيه قيم Y مع زيادة قيم X ، ونستنتج منه وجود علاقة خطية طردية بين المتغيرين (X, Y) .

شكل الانتشار (٥ - ٣)

تكون فيه النقاط منتشرة حول خط مستقيم وفيه تنقص قيم Y مع زيادة قيم X ونستنتج منه وجود علاقة خطية عكسية بين المتغيرين (X, Y) .

شكل الانتشار (٥ - ٤)

تكون فيه النقاط منتشرة حول منحنى نستنتج منه وجود علاقة غير خطية بين المتغيرين (X, Y) .

وسوف نكتفي في هذا الفصل بإيجاد مقاييس تقيس قوة الارتباط بين المتغيرين (X, Y) في الحالة الخطية فقط وسندرس منها معامل الارتباط الخطي لبيرسون (Person) ومعامل ارتباط الرتب لسبيرمان (Spearman) ، كما سوف ندرس معامل الاقتران ومعامل التوافق لكرامير (Cramear) ، وكذلك دراسة معادلة خط الانحدار للمتغير Y على المتغير X أو العكس للمتغير X على Y .

(٥ ـ ٢) معامل الارتباط الخطي لبيرسون

(Coefficient of linear correlation)

ويستخدم معامل الارتباط الخطي لبيرسون لقياس التغير الذي يطرأ على المتغير Y عندما تتغير قيم X أو العكس. ويستخدم عادة في حالة البيانات الكمية وسنوضح طريقة حسابه في حالتين :

١) البيانات المباشرة (عدد الأفراد n صغير نسبيًّا) .

٢) البيانات المبوبة (عدد الأفراد n كبير نسبيًّا) .

وسنتناول كل حالة بالتفصيل كما يلي :

(٥ ـ ٢ ـ ١) معامل ارتباط بيرسون للبيانات المباشرة

إذا كان لدينا ازدواج المشاهدات التالية :

$$(x_1, y_1), (x_2, y_2), \ldots\ldots , (x_n, y_n)$$

فإن معامل الارتباط r لبيرسون يعطى بمتوسط مجموعة حاصل ضرب القيم المعيارية للمتغيرين X', Y' ويكون كالتالي :

$$r = \frac{1}{n} \sum x'y' \quad \ldots\ldots\ldots\ldots \quad (٥ ـ ١)$$

حيث

$$y' = \frac{y - \bar{y}}{\sigma_y} \ , x' = \frac{x - \bar{x}}{\sigma_x}$$

أي أن العلاقة (٥ ـ ١) يمكن كتابتها كالتالي :

$$r = \frac{1}{n} \frac{\sum (x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y} \quad \text{.....................} \quad (٥ - ٢)$$

ويفضل في حالة البيانات المأخوذة من عينة وضع العلاقة (٥ ـ ٢) في الصورة التالية :

$$r = \frac{\sum (x-\bar{x})(y-\bar{y})}{(n-1) S_x S_y} \quad \text{.....................} \quad (٥ - ٣)$$

ويكون لمعامل الارتباط (r) الخصائص التالية .

(٥ ـ ٢ ـ ٢) خصائص معامل الارتباط الخطي لبيرسون

١) قيمته تساوي صفرًا عندما تكون الظاهرتان مستقلتين تمامًا ولا يوجد ارتباط بينهما ألبتة .

٢) قيمته مقدار موجب عندما يكون الارتباط بين المتغيرين طرديًا . ويكون قويًا عندما يكون المقدار الموجب قريبًا من الواحد الصحيح وضعيفًا عندما يكون المقدار الموجب قريبًا من الصفر.

٣) قيمته مقدار سالب عندما يكون الارتباط بين المتغيرين عكسيًا، ويكون قويًا عندما تقترب قيمته من المقدار ـ ١ وضعيفًا عندما تقترب من الصفر.

ملاحظة (١)

مستوى هذا الكتاب لا يتسع لعرض النظريات المتطلبة لإجراء معنوية الارتباط والتي سندرسها في كتاب المستوى الثاني .

ملاحظة (٢)

ويجب ملاحظة عدم ربط الارتباط بين متغيرين بالسببية أي أن التغير في أحد المتغيرين ليس بالضرورة يؤدي إلى التغير في المتغير الثاني (يتسبب فيه) فمثلاً إذا وجدنا ارتباطًا قويًا بين سلسلة من بيانات خاصة بأعداد الطلبة في سنوات متتالية وسلسلة بيانات الزيادة في عدد حالات الزواج أو العكس بالعكس . أو إذا كان يوجد ارتباط عكسي قوي بين سلسلة مبيعات البطاطين في المملكة العربية السعودية في شهور سنة

ما مع درجات الحرارة في انجلترا في نفس شهور السنة فلا يمكن تفسير ذلك بأن انخفاض درجة الحرارة في انجلترا يتسبب في زيادة مبيعات البطاطين في السعودية.

ملاحظـة

المعادلة (٥ ـ ٣) السابقة لمعامل الارتباط r تكون عبارة عن مقياس نسبي يعتمد على قياس مقدار انحراف المتغيرين (X, Y) عن وسطيهما الحسابي $(\bar{X}, \bar{Y})$ وبذلك فإن إضافة أو طرح مقدار ثابت من كليهما لا يؤثر على قيمة r وتبقى كما هي ويمكن إثبات ذلك كما يلي:

إذا كان لدينا أزواج القيم التالية:

$$(x_1, y_1)\ ,\ (x_2, y_2)\ ,\ \ldots\ ,\ (x_n, y_n)$$

فإن انحرافات القيم لـ X من قيمة ثابتة A تكون كالتالي:

$$d_1 = x_1 - A\ \ ,\ \ d_2 = x_2 - A\ \ ,\ \ \ldots\ldots\ \ ,\ \ d_n = x_n - A$$

وأن انحرافات القيم لـ Y من قيمة ثابتة B تكون كالتالي:

$$d_1' = y_1 - B\ \ ,\ \ d_2' = y_2 - B\ \ ,\ \ \ldots\ldots\ \ ,\ \ d_n' = y_n - B$$

وعليه فإن معامل الارتباط $r_{(d.d')}$ لأزواج انحرافات القيم التالية:

$$(d_1, d_1')\ ,\ (d_2\ , d_2')\ ,\ \ldots\ldots\ ,\ (d_n\ , d_n')$$

يكون

$$r_{(d,d')} = \frac{\sum (d - \bar{d})(d' - \bar{d}')}{(n-1)\, S_d\, S_{d'}}$$

وبالتعويض عن d و d' بالقيم $(x-A), (y-B)$ فنحصل على

$$r_{(d,d')} = \frac{\sum (x - A - \bar{x} + A)(y - B - \bar{y} + B)}{(n-1)\, S_x\, S_y}$$

$$r_{(d,d')} = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n-1)\, S_x\, S_y}$$

$$r_{(d,d')} = r_{(x,y)}$$

وبـالمثـل يمكن إثبات أنه إذا ضربنا أو قسمنا قيها لمتغيرين x, y على مقدارين ثابتين B و A على التوالي فإن قيمة معامل الارتباط $r_{(x,y)}$ تظل كما هي ولا تتغير (برهن ذلك). وحسـاب معـامل الارتباط من المعادلة (٥ ـ ٣) يتطلب حساب كل من قيم S_x و S_y و $\bar{x}$ و $\bar{y}$ مما يجعل الحسابات طويلة وعرضة للخطأ ولذلك أمكن تبسيط المعادلة (٥ ـ ٣) وذلك بالتعويض من هذه القيم من التعاريف السابقة لها فيصبح معامل الارتباط (r) كالتالي:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)\,(n\sum y^2 - (\sum y)^2)}} \quad \text{......................} \quad (٥ - ٤)$$

وواضح من المعادلة (٥ ـ ٤) أنه لحساب r يتطلب منا إيجاد قيم المقادير $\sum x, \sum y, \sum xy, \sum x^2, \sum y^2$ فقط ونوضح ذلك بالمثال التالي:

مثال (٥ ـ ١)

الجـدول التالي يبين درجات مجموعة مكونة من ثمانية طلّاب في كل من مادتي الإحصاء والرياضيات في أحد الامتحانات للأعمال الفصلية.

هل هناك علاقة بين تحصيل الطلاب في المادتين؟

الإحصاء Y	13	9	19	15	11	8	16	11
الرياضيات X	15	7	17	15	10	9	14	10

الحـــل

لتبسيط البيانات التي بالجدول نطرح مقدارًا ثابتًا $a=10$ من كل من قيم X وقيم Y

ولتبسيط الحل نكون الجدول التالي:

x	y	x′ = x − 10	y′ = y − 10	x′ y′	x'^2	y'^2
13	15	3	5	15	9	25
9	7	−1	−3	3	1	9
19	17	9	7	63	81	49
15	15	5	5	25	25	25
11	10	1	0	0	1	0
8	9	−2	−1	2	4	1
16	14	6	4	24	36	16
11	10	1	0	0	1	0
Σ		22	17	132	158	125

$$r = \frac{n\sum x' y' - \sum x' \sum y'}{\sqrt{(n\sum x'^2 - (\sum x')^2)\,(n\sum y'^2 - (\sum y')^2)}}$$

$$r = \frac{8 \times 132 - 22 \times 17}{\sqrt{[8 \times 158 - (22)^2]\,[8 \times 125 - (17)^2]}}$$

$$r = \frac{690}{\sqrt{780 \times 711}} = \frac{690}{744.7} = 0.93$$

أي يوجد ارتباط قوي جدًّا بين درجات تحصيل الطلاّب في المادتين.

(٥ ـ ٢ ـ ٣) معامل الارتباط الخطي لبيرسون في حالة البيانات المبوبة

في حالة البيانات المبوبة تصبح المعادلة (٥ ـ ٤) لمعامل الارتباط في الصورة التالية:

$$r = \frac{n\sum f_{xy}xy - (\sum f_x x)(\sum f_y y)}{\sqrt{[n\sum f_x x^2 - (\sum f_x x)^2 [n\sum f_y y^2 - (\sum f_y y)^2]}} \quad \text{......................} \quad (٥ ـ ٥)$$

ونوضح استخدام العلاقة (٥ ـ ٥) في حساب معامل الارتباط (r) للبيانات المبوبة بالمثال التالي.

مثال (٥ ـ ٢)

أوجد معامل الارتباط r لدرجات الطلاب في كل من مادتي الإحصاء والرياضيات والمبين في مثال (٢ ـ ٣) في الفصل الثاني.

الحـــل

ولتبسيط الحل نوضح خطوات الحل في الجدول التالي حيث يمكن استخدام الطريقة المختصرة في الحساب وهي أن نطرح وسطًا فرضيًّا (a) ونقسم على طول الفئة (10) حيث الفئات منتظمة ومغلقة من الطرفين لهذا المثال. وفي هذا المثال يكون الوسط الفرضي للمتغيرين (X, Y) هو مركز الفئة (79 – 70) أي $a = 74.5$ حيث يناظر كل منهما أكبر تكرار ونقسم على طول الفئة لكل منهما (L = 10) حيث كل منهما متساويا الطول وعليه فيكون لدينا الانحرافات $d_X = \frac{X-74.5}{10}$ بالنسبة للمتغير X درجات الإحصاء، $d_Y = \frac{Y-74.5}{10}$ بالنسبة للمتغير Y درجات الرياضيات. f_X ترمز إلى تكرار فئات المتغير X (مجموع التكرار لكل عمود أفقي) و f_Y ترمز إلى تكرار فئات المتغير Y (مجموع التكرار لكل عمود رأسي) و f_{XY} ترمز إلى التكرار المشترك لفئات المتغيرين (X,Y) الموجود داخل المربعات بالجدول. وتوضح طريقة الحساب بالجدول التالي.

الرياضيات Y الاحصاء X	50-59	60-69	70-79	80-89	90-99	f_X	d_X	f_Xd_X	$f_Xd_X^2$	$f_{XY}d_Xd_Y$
50 – 59	3 $\boxed{12}$					3	−2	−6	12	12
60 – 69	1 $\boxed{2}$	4 $\boxed{4}$				5	−1	−5	5	6
70 – 79		2 $\boxed{0}$	8 $\boxed{0}$			10	0	0	0	0
80 – 89			1 $\boxed{0}$	6 $\boxed{6}$	1 $\boxed{2}$	8	1	8	8	8
90 – 99					4 $\boxed{16}$	4	2	8	16	16
f_Y	4	6	9	6	5	30		5	41	42
d_Y	−2	−1	0	1	2					
f_Yd_Y	−8	−6	0	6	10	2				
$f_Yd_Y^2$	16	6	0	6	20	48				
$f_{XY}d_Xd_Y$	14	4	0	6	18	42				

المقادير الموجودة في أركان المربعات هي : $f_{XY}d_Xd_Y$ وهي عبارة عن حاصل ضرب التكرار الموجود داخل المربع مضروبًا في كل من مراكز الفئات المختصرة d_X, d_Y المناظرة لنهاية الصف والعمود لهذا المربع ثم نجمع القيم في الأركان أفقيًا مرة ورأسيًا مرة أخرى فنحصل على مجموع متساوٍ لكل منهما وهو $\sum f_{XY}d_Xd_Y$ كما هو موضح في الجدول السابق ثم نحسب r كالتالي :

$$r = \frac{n\sum f_{XY}d_Xd_Y - (\sum f_Xd_x)(\sum f_Yd_y)}{\sqrt{[n\sum f_Xd_x^2 - (\sum f_Xd_x)^2]\,[n\sum f_Yd_y{}^2 - (\sum f_Yd_y)^2]}} \quad \text{......}\quad (٥ \text{ - } ٥)$$

ونعوض عن القيم السابقة من الجدول السابق كالتالي :

$$r = \frac{30 \times 42 - (2)(5)}{\sqrt{[30 \times 48 - (2)^2][30 \times 41 - (5)^2]}}$$

$$r = \frac{1250}{\sqrt{1346 \times 1235}} = \frac{1250}{1331.71} = 0.939$$

أي يوجد ارتباط طردي قوي لدرجات الإحصاء والرياضيات.

(٥ ـ ٣) معامل الارتباط للرتب لسبيرمان (Spearman)

معامل الارتباط الخطي لبيرسون الذي سبق الكلام عنه يقيس مقدار الارتباط بين متغيرين وذلك في حالة البيانات الكمية فقط وفي بعض الأحيان يكون مطلوبًا إيجاد الارتباط بين متغيرين على صورة بيانات وصفية يمكن وضعها في صورة ترتيبية مثل تقديرات الطلاب في مادتين مختلفتين فتكون التقديرات هي E, D, C, B, A فإنه يصعب حساب معامل الارتباط بطريقة بيرسون لذلك نشأت الحاجة إلى إيجاد مقياس يعطي قوة الارتباط للبيانات الوصفية. وهذا المقياس هو ما يسمى بمعامل ارتباط الرتب لسبيرمان وهو يعطي مقياسًا للارتباط في كل من البيانات الكمية والوصفية التي لها صفة الترتيب مثل تقديرات الطلاب. فإنه يمكن إعطاء رتب لها من حيث كبر التقدير وصغره. وكذلك البيانات الكمية نلاحظ أن رتب المتغيرين (X, Y) تزيد وتنقص حسب زيادة ونقص كل من قيم المتغيرين (X, Y) ولذلك فإن حساب معامل الارتباط للرتب يقترب كثيرًا من معامل ارتباط بيرسون ولكن يمتاز عنه في السهولة والدقة وخاصة عندما تكون أزواج القيم أقل من (n=30) ويعطى معامل ارتباط الرتب بالعلاقة التالية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \qquad \text{.............................} \qquad (٥ - ٦)$$

حيث r_s معامل ارتباط الرتب لسبيرمان لـ n من أزواج القيم (X, Y)، d هي الفرق بين رتب أزواج القيم (X, Y) ويمكن توضيح حساب الرتب كالتالي:

المقصود بالرتب هنا هو إيجاد رتب القراءات (X, Y) مع بقاء كل قراءة مكانها وذلك بأن نتصور ترتيب البيانات تصاعديًا أو تنازليًا. ففي حالة الترتيب التصاعدي تكون أصغر قراءة رتبتها تساوي واحدًا والقيمة التي تليها في الكبر تأخذ رتبة 2 وهكذا وفي حالة تساوي قيمتين نعطي لكل قيمة منهما رتبة تساوي الوسط الحسابي لرتبتيهما وسوف نوضح ذلك بالأمثلة التالية.

مثال (٥ ـ ٣)

أوجد رتب X التي قيمها معطاة في الجدول التالي:

x	10	4	5	7	2

الحــل

نتصور ترتيب قيم X تصاعديًا فتكون القيمة 2 رتبتها 1 والقيمة 4 رتبتها 2 والقيمة 5 رتبتها 3 وهكذا يمكن تلخيص ذلك في الجدول التالي:

x	10	4	5	7	2
رتبة x	5	2	3	4	1

مثال (٥ ـ ٤)

أوجد رتب التقديرات الآتية:

B, C, B, E, D, D, A

الحــل

إذا تصورنا ترتيب البيانات تصاعديًا فإننا نجد E تكون رتبتها 1 ، D تأخذ الرتبتين 2,3 وعليه فإن كل قيمة من D تأخذ الرتبة $\frac{2+3}{2}$ أي رتبة كل من D هي 2.5، ورتبة C تكون 4 ، B رتبتها $\frac{5+6}{2}$ ، أي 5.5 ، ورتبة A تكون 7 .

يمكن تلخيص ذلك في الجدول التالي:

B	C	B	E	D	D	A	التقديرات X
5.5	4	5.5	1	2.5	2.5	7	رتـــب X

مثال (٥ ـ ٥)

أوجد معامل ارتباط الرتب لدرجات الطلاّب للأعمال الفصلية لمادتي الإحصاء والرياضيات في مثال (٥ ـ ١) السابق.

الحـــل

نلخص الحساب والحل في الجدول التالي:

$d_2 = (a-b)^2$	$d = a-b$	رتبة Y b	رتبة X a	درجة الرياضيات	درجة الإحصاء
2.25	−1.5	6.5	5	15	13
1	1	1	2	7	9
0	0	8	8	17	19
0.25	−0.5	6.5	6	15	15
0	0	3.5	3.5	10	11
1	−1	2	1	9	8
4	2	5	7	14	16
0	0	3.5	3.5	10	11
$\sum d^2 = 8.5$					

$$\because\ r_s = 1 - \frac{6 \sum d^2}{n\,(n^2-1)}$$

$$\therefore\ r_s = 1 - \frac{6 \times 8.5}{8\,(64-1)} = 1 - \frac{51}{504} = 0.9$$

أي أن الارتباط بين درجات المادتين يكون ارتباطًا قويًا كما سبق حسابه في معامل ارتباط بيرسون .

مثال (٥ ـ ٦)

أوجد معامل ارتباط الرتب لتقديرات الطلاب في مادتي الإحصاء والرياضيات كما هو موضح بالجدول التالي :

الرياضيات X	A	C	C	C	B	D
الإحصاء Y	B	B	D	C	A	E

الحـــل

الرياضيات X	الإحصاء Y	رتبة X a	رتبة Y b	$d = a - b$	$d^2 = (a-b)^2$
A	B	6	4.5	1.5	2.25
C	B	3	4.5	− 1.5	2.25
C	D	3	2	1	1
C	C	3	3	0	0
B	A	5	6	− 1	1
D	E	1	1	0	0
					6.5

$$\because\ r_s = 1 - \frac{6 \sum d_2}{n(n^2 - 1)}$$

$$\therefore\ r_s = 1 - \frac{6 \times 6.5}{6(36-1)} = 1 - 0.186 = 0.814$$

أي أنه يوجد ارتباط طردي قوي بين تقديرات الرياضيات والإحصاء.

(٥ ـ ٤) معامل الاقتران ومعامل التوافق (Coefficient of contingency)

لقد سبق أن وضحنا بأن معامل ارتباط بيرسون يعطي قوة الارتباط في حالة البيانات الكمية سواء كانت مبوبة أو غير مبوبة وكذلك معامل ارتباط الرتب لسبيرمان وهو يستخدم لايجاد قوة الارتباط للرتب في حالة البيانات الكمية والوصفية التي لها صفة الترتيب ولكن قد تكون هناك بيانات وصفية لها صفات مميزة ولكن لا نستطيع ترتيبها مثل الحالة الاجتماعية (أعزب ـ متزوج ـ أرمل ـ مطلق) وكذلك لون البشرة (أبيض ـ أسود ـ . . . . .) وكذلك في حالة الجنسية (سعودي ـ مصري ـ سوري ـ . . . الخ). ولقياس قوة الارتباط لهذه البيانات نشأت الحاجة إلى إيجاد مقاييس تقيس الارتباط بين هذه الصفات وسنذكر منها معامل الاقتران ويستخدم عندما يكون لكل من الظاهرتين صفتين فقط فعند قياس قوة الارتباط بين التدخين والتعليم فإننا نجد التعليم ينقسم إلى صفتين (متعلم ـ غير متعلم) والتدخين ينقسم إلى صفتين (يدخن ـ ولا يدخن) وكذلك دراسة معامل التوافق عندما تكون كل من الظاهرتين أو إحداهما لها أكثر من صفتين كما سنوضح ذلك بالتفصيل فيما يلي.

(٥ ـ ٤ ـ ١) معامل الاقتران

يستخدم معامل الاقتران لقياس قوة الارتباط بين ظاهرتين كل ظاهرة منهما ذات صفتين فقط وسوف يرمز له بالرمز c.c مثل دراسة علاقة قوة الارتباط بين التدخين والتعليم والجدول التالي يبين التكرارات للصفات.

فيكون معامل الاقتران c.c كالتالي:

$$c.c = \frac{AD - BC}{AD + BC} \qquad \text{.........................} \qquad (٥ - ٧)$$

ونوضح ذلك بالمثال التالي .

(1)

التدخين / التعليم	يـدخــن	لا يدخــن
متعلـم	A	B
غير متعلم	C	D

(2)

مثال (٥ ـ ٧)

عند دراسة علاقة التدخين بالتعليم في إحدى المؤسسات أخذت عينة مكوّنة من 17 شخصًا وكانت النتائج موضحة بالجدول الآتي :

التدخين / التعليم	يـدخــن	لا يدخــن
متعلـم	5	5
غير متعلم	3	4

احسب معامل الاقتران c.c بين التدخين والتعليم .

الحــل

$$\because \; c.c = \frac{AD - CB}{AD + CB}$$

ونوضح ذلك بالمثال التالي .

$$\therefore\ c.c = \frac{5x4 - 3x5}{5x4 + 3x5} = \frac{5}{35} = 0.14$$

وهو ارتباط ضعيف.

(٥ ـ ٤ ـ ٢) معامل التوافق

إذا كانت بيانات الظاهرتين التي لدينا عبارة عن بيانات وصفيّة لكل منهما أو وصفية لإحداهما وكميّة للأخرى وكانت مقسمة إلى أكثر من نوعين (أي أن الجدول يحتوي على أكثر من أربع خانات) فإن معامل الاقتران السابق لا يصلح في هذه الحالة ونستخدم مقياسًا آخر هو معامل التوافق c.c (أوجده كرامير Cramer 1946) ولحساب معامل التوافق نفرض أنه لدينا الظاهرة (X) لها r من الصفات والظاهرة الثانية (Y) لها s من الصفات ونوضح جدول الاقتران بين الظاهرتين كما يلي:

الصفة Y / الصفة X	y_1	y_2	....	y_s	المجموع
x_1	f_{11}	f_{12}	....	f_{1s}	$f_{1.}$
x_2	f_{21}	f_{22}	....	f_{2s}	$f_{2.}$
⋮	⋮	⋮		⋮	⋮
x_r	f_{r1}	f_{r2}	....	f_{rs}	$f_{r.}$
المجموع	$f_{.1}$	$f_{.2}$	....	$f_{.s}$	$f_{..}$

نحسب المقدار B ومنه نحسب معامل التوافق c.c بالعلاقة التالية:

$$c.c = \sqrt{\frac{B-1}{B}} \qquad \text{.........} \qquad (٥ ـ ٨)$$

حيث

$$B = \frac{(f_{11})^2}{f_{.1} \times f_{1.}} + \frac{(f_{12})^2}{f_{.2} \times f_{1.}} + \dots + \frac{(f_{rs})^2}{f_{.s} \times f_{r.}}$$

ونوضح ذلك بالمثال التالي.

مثال (٥ ـ ٨)

عند دراسة العلاقة بين الرائحة ولون الزهور لعينة مكونة من 30 زهرة كانت لدينا النتائج التالية:

الرائحة Y / اللون X	بدون رائحة	له رائحة	المجموع
أصفر	6	4	10
أبيض	7	2	9
أحمر	6	5	11
المجموع	19	11	30

أحسب معامل التوافق c.c بينا للون والرئحة للزهور.

الحل

نحسب قيمة B كالتالي:

$$B = \frac{(6)^2}{19 \times 10} + \frac{(7)^2}{19 \times 9} + \frac{(6)^2}{19 \times 11} + \frac{(4)^2}{11 \times 10} + \frac{(2)^2}{11 \times 9} + \frac{(5)^2}{11 \times 11}$$

$$= 0.19 + 0.29 + 0.17 + 0.15 + 0.04 + 0.21 = 1.05$$

ويكون معامل التوافق c.c كالتالي:

$$\because\ c.c = \sqrt{\frac{B-1}{B}}$$

$$\therefore\ c.c = \sqrt{\frac{1.05-1}{1.05}} = \sqrt{\frac{0.05}{1.05}} = 0.22$$

ونلاحظ أن قيمة معامل التوافق c.c = 0.22 تبين مقدار قوة الارتباط وهي ضعيفة في هذا المثال.

(٥ ـ ٤ ـ ٣) معامل ارتباط كندال

قدم كندال عام ١٩٣٨م معامل ارتباط لقياس ارتباط متغيرين y, x على سبيل المثال وذلك باستخدام رتب هذين المتغيرين. ويرمز لمعامل ارتباط كندال بالرمز τ ويقرأ (تو)، وتعطى صيغته كما يلي:

$$\tau = 1 - \frac{4Q}{n(n-1)} \qquad (٥ - ٩)$$

حيث n هو عدد أزواج القيم (x, y) في الدراسة، وفيما يلي نقدم شرحًا مبسطًا لطريقة حساب قيمة Q:

ا ـ صنِّف رتب المتغير x تصاعديًا مع نقل رتب المتغير y المناظرة لكل رتبة من رتب x تبعًا لذلك.

ب ـ بدءًا من القيمة الأصغر لرتب x ، نحسب لكل رتبة من رتب y عدد الرتب التي عن يمينها وأصغر منها ونقوم بتسجيلها. ونكرر ذلك لجميع رتب y ، وتكون قيمة Q هي مجموع عدد الرتب التي تم تسجيلها. ونوضح ذلك بالمثال التالي.

مثال (٥ ـ ٩)

قام اثنان من المدربين بإعطاء أفراد فريق رياضي مكوّن من سبعة أشخاص رتبًا حسب أفضلية كل لاعب في نظر كل منهما كما هو موضح بالجدول التالي:

رقم اللاعب	1	2	3	4	5	6	7
x رتب المدرب الأول	4	1	6	5	3	2	7
y رتب المدرب الثاني	4	2	5	6	1	3	7

احسب معامل ارتباط كندال بين رتب المدربين.

الحـــل

نضيف رتب المدرب الأول (x) تصاعديًّا مع نقل كل رتبة من رتب المدرب الثاني (y) المناظرة لكل منهما تبعًا لذلك، كما هو موضح بالجدول التالي :

رقـم اللاعـب	2	6	5	1	4	3	7
رتب المدرب الأول مرتبة تصاعديًّا	1	2	3	4	5	6	7
رتب المدرب الثاني المناظرة لها	2	3	1	4	6	5	7
عدد الرتب التي عن يمين كل من y وأقل منها	1	1	0	0	1	0	0

وتم الحصول على الصف الأخير من الجدول السابق كما يلي :
الرتبة الأولى لـ y هي 2 والرتبة الثانية لـ y هي 3 ويقع إلى يمين كل منهما رتبة واحدة فقط هي 1. أما الرتبة الثالثة لـ y هي 1 ولا يقع على يمينها أي رتبة أصغر منها فتكون القيمة التي تضاف هي صفر، وبنفس الطريقة يمكن إكمال الصف الأخير من الجدول السابق وبذلك تكون Q كما يلي :

$$Q = 1 + 1 + 0 + 0 + 1 + 0 = 3 \qquad (١٠ - ٥)$$

وبالتعويض عن قيمة Q من (٥ - ١٠) في العلاقة (٥ - ٩) نحصل على معامل ارتباط كندال τ كما يلي :

$$\tau = 1 - \frac{4(3)}{7(7-1)} = 0.714$$

حيث $n = 7$ لاعبين ومن القيمة $\tau = 0.714$ نلاحظ أن معامل ارتباط كندال لرتب المدربين هو ارتباط قوي نوعًا ما .

(٥ - ٥) خط الانحدار (Regression line)

سبق لنا أن بَيَّنّا أنه إذا كان لدينا متغيران (X,Y) أحدهما متغير مستقل (X) مثلاً والآخ رمتغير تابع وليكن (Y) مثلاً لظاهرتين محل الدراسة مثل العلاقة بين الطول (X) والـوزن (Y) . أو العلاقة بين الإنفاق (Y) والدخل (X) . فإنه يمكن عمل أشكال الانتشار لهذه المتغيرات. وقد تم دراسة إيجاد معامل الارتباط بعدة طرق لقياس قوة الارتباط بين المتغيرين (X,Y) وذلك في الحالة التي يكون فيها شكل الانتشار في صورة

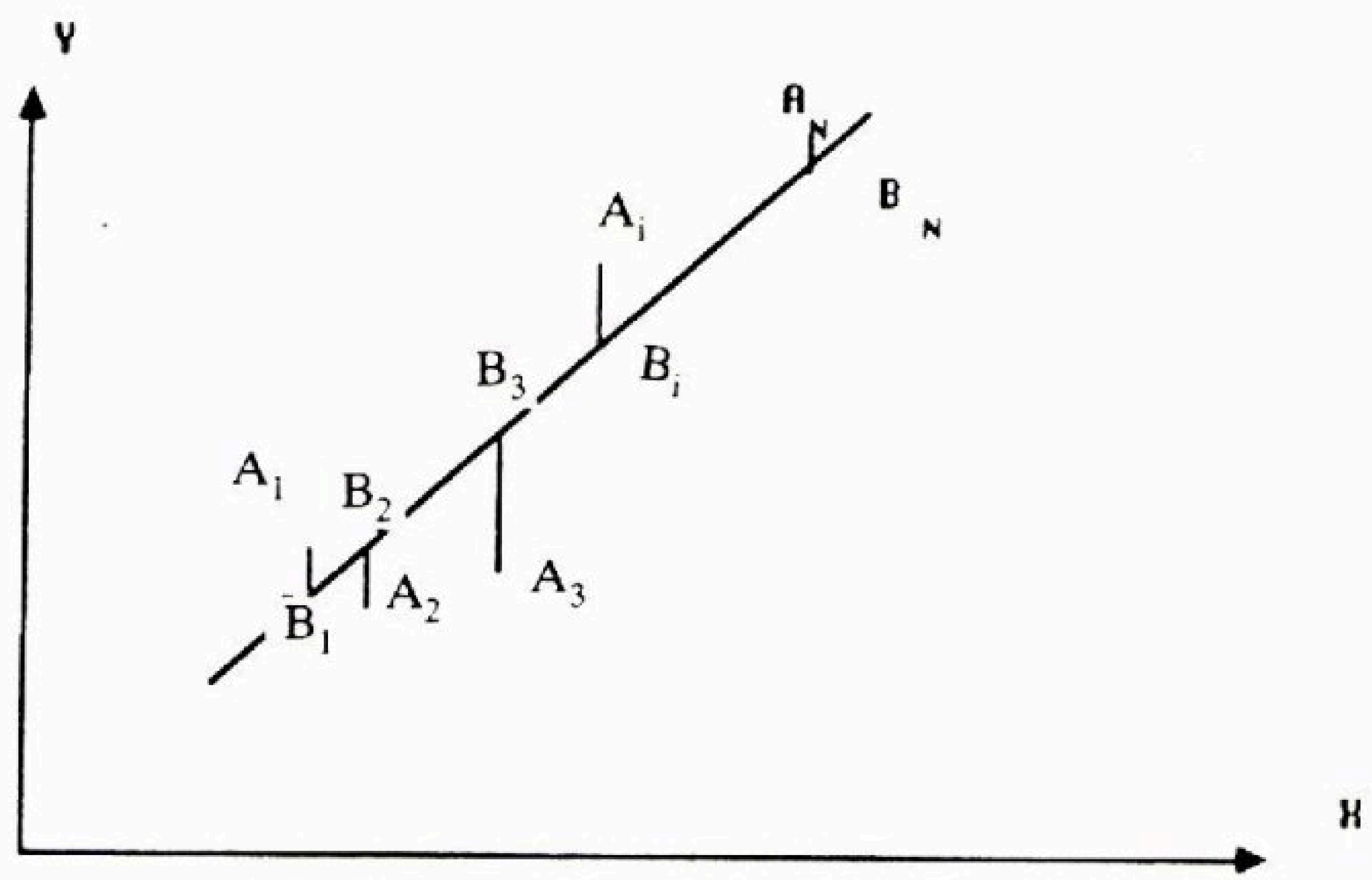
خطية . وفيما يلي نبحث عن إيجاد معادلة رياضية تمثل أحسن توفيق لخط مستقيم يعبر عن البيانات في شكلها الخطي . وهذه العلاقة تسمى بمعادلة خط الانحدار. فإذا كان X متغيراً مستقلاً، و Y متغيراً تابعًا، فإن المعادلة التي نحصل عليها تسمى بمعادلة خط انحدار Y على X وهي على الصورة التالية :

$$y = mx + c \qquad \text{(٥ ـ ٩)}$$

وإذا كان المتغير Y متغيراً مستقلاً والمتغير X متغيراً تابعًا فإن معادلة الخط المستقيم تسمى بمعادلة خط انحدار X على Y وتعطى بالعلاقة التالية :

$$x = m'y + c' \qquad \text{(٥ ـ ١٠)}$$

والغرض من إيجاد معادلة خط الانحدار هو التنبؤ بقيمة المتغير التابع لقيمة محددة من قيم المتغير المستقل. ويمكن الحصول على رسم خط الانحدار للبيانات من شكل الانتشار وذلك بالتمهيد باليد، ولكن هذا التمهيد باليد يختلف من شخص إلى آخر ولذلك دعت الحاجة إلى إيجاد خط الانحدار بطريقة لا تعتمد على الرسم وإنما تعتمد على استخدام البيانات المعطاة للحصول على قيم الثوابت m,c في حالة خط انحدار Y على X وللحصول على m′,c′ في حالة خط انحدار X على Y وفي حالة انحدار Y على X يمكن إيجاد قيمة الثابتين m,c بطريقة المربعات الصغرى كما يلي :



شكل (٥ ـ ٥) يبين انحرافات النقط التي تمثل المشاهدات عن خط الانحدار

X Y x y $y = mx + c$

A_1 B_2 B_3 B_4 B_i B_n

A_1 A_2 A_3 A_4 A_i A_n

نفرض أن لدينا مجموعة أزواج المشاهدات (x_1, y_1) , ... , (x_n, y_n) ويرسم شكل الانتشار لهذه الأزواج نحصل على النقاط A_1 , ... , A_n نفرض أن خط الانحدار $y = mx + c$ كما هو موضح بالرسم. وبرسم خطوط موازية لمحور y مارة بالنقاط A_1 , ... , A_n فإنها تقطع خط الانحدار في النقاط B_1 , ... , B_n . نرمز للأطوال A_1B_1 , ... , A_nB_n بالقيم D_1 , ... , D_n وهذه الأطوال عبارة عن انحرافات المشاهدات عن خط الانحدار. وقد تكون هذه الانحرافات سالبة أو موجبة حسب وضع النقط A_1 , ... , A_n بالنسبة لخط الانحدار.

لكي نحصل على أجود خط لا بد أن يكون مجموع مربعات هذه الانحرافات أقل ما يمكن أي أن:

$$D = \sum_{i=1}^{n} (mx_i + c - y_i)^2$$

أقل ما يمكن.

بتفاضل D بالنسبة إلى m ، c ومساواة الناتج بالصفر نحصل على:

$$\sum y = m \sum x + nc \qquad \text{.................} \qquad (٥ - ١١)$$

$$\sum xy = m \sum x^2 + c \sum x \qquad \text{.................} \qquad (٥ - ١٢)$$

وبحل هاتين المعادلتين نحصل على:

$$m = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$\text{.................} \qquad (٥ - ١٣)$$

$$c = \frac{\sum y}{n} - m\,\frac{\sum x}{n}$$

وبالطريقة نفسها يمكن الحصول على m', c' من المعادلة (٥ ـ ١٠) فنحصل على:

$$m' = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$$

(٥ ـ ١٤)

$$c' = \frac{\sum x}{n} - m' \frac{\sum y}{n}$$

مثال (٥ ـ ٩)

أوجد معادلة خط انحدار درجات الإحصاء (Y) على درجات الرياضيات (X) في مثال (٥ ـ ١).

الحـــل

نلخص الحسابات في الجدول التالي:

x	y	xy	x^2	y^2
15	13	159	225	169
7	9	63	49	81
17	19	323	289	361
15	15	225	225	225
10	11	110	100	121
9	8	72	81	64
14	16	224	196	256
10	11	110	100	121
$\sum$ 97	102	1322	1265	1398

يكون عدد أزواج القيم $n = 8$

$$\because\ m = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$\therefore\ m = \frac{8 \times 1322 - 97 \times 102}{8 \times 1265 - (97)^2}$$

$$= \frac{10576 - 97 \times 102}{10120 - 9409} = \frac{682}{711} = 0.96$$

$$\because\ c = \frac{\sum y}{n} - m.\ \frac{\sum x}{n}$$

$$\therefore\ c = \frac{102}{8} - 0.96\,(\frac{97}{8})$$

$$= 12.75 - 11.64 = 1.11$$

أي أن معادلة خط انحدار Y على X هي :

$$Y = 0.96\,X + 1.11$$

وكذلك يمكن حساب معادلة انحدار X على Y كالتالي :

$$\because\ m' = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$$

$$\therefore\ m' = \frac{8 \times 1322 - 97 \times 102}{8 \times 1398 - (102)^2}$$

$$= \frac{682}{780} = 0.87$$

$$\because\ c' = \frac{\sum x}{n} - m'\ \frac{\sum y}{n}$$

$$\therefore\ c' = \frac{97}{8} - 0.87\,(\frac{102}{8})$$

$$= 12.125 - 11.09 = 1.035$$

أي أن معادلة خط انحدار X على Y هي :

$$X = 0.87Y + 1.035$$

(٥ ـ ٦) تمـارين

(١) في أحد أماكن بيع السيارات كانت المبيعات كالتالي:

عمر السيارة بالسنوات X	3	2	1	1	5	6	1	4
ثمن البيع بمئات الريالات Y	31	44	60	70	18	17	71	29

ا) أوجد معامل الارتباط بين عمر السيارة وثمن البيع بطريقة بيرسون.

ب) أوجد خط انحدار Y على X .

جـ) أوجد خط انحدار X على Y .

(٢) من البيانات التالية:

x	0	1	2	3	4	5	2
y	-1	2	2	8	4	14	5

ا) أوجد معامل الارتباط بين المتغيرين Y و X بطريقة بيرسون وسيبرمان.

ب) أوجد خط انحدار Y على X .

(٣) البيانات التالية تمثل تقديرات ثمانية طلاب في مادتي الكيمياء والفيزياء.

الكيمياء	A	B	D	E	C	D	E	B
الفيزياء	A	C	E	D	C	D	E	B

أوجد معامل الارتباط لتقديرات الكيمياء والفيزياء.

(٤) الجدول التالي يمثل الدخل X والإنفاق Y من الأسر بمئات الريالات.

x	56	66	42	44	38	27	39	40
y	31	38	27	22	19	25	20	28

ا) أوجد معامل ارتباط بيرسون وسبيرمان للدخل والإنفاق.

ب) أوجد خط انحدار Y على X .

جـ) أوجد خط انحدار X على Y ،

د) أوجد قيمة الإنفاق عندما يصبح الدخل 6000 ريال.

(٥) الجدول التالي يمثل درجات الحرارة والمبيعات في إحدى محطات المحروقات.

درجة الحرارة X	25	30	32	33	35	40	37
المبيعات بمئات الريالات Y	44	45	33	38	30	27	41

ا) أوجد معامل ارتباط درجة الحرارة والمبيعات بطريقتين مختلفتين.

ب) أوجد خط انحدار Y على X .

جـ) أوجد خط انحدار X على Y .

د) أوجد قيمة المبيعات عندما تكون درجة الحرارة 45° .

(٦) الجدول التالي يوضح السن X وضغط الدم Y لثمان من الإناث:

السن X	42	36	63	55	42	60	49	68
ضغط الدم Y	125	118	140	150	140	155	145	152

ا) أوجد معامل الارتباط بين Y و X .

ب) أوجد خط انحدار Y على X .

جـ) أوجد مقدار ضغط الدم لامرأة عمرها 46 سنة .

(٧) الجدول التالي يمثل أوزانًا لعينة مكونة من ٨ أشخاص X وأكبر الأبناء Y .

68	64	71	67	70	62	62	64	أوزان الآباء X
69	65	70	67	69	65	67	66	أوزان الأبناء Y

ا) أوجد معامل الارتباط بين Y و X .

ب) أوجد خط انحدار Y على X .

جـ) أوجد خط انحدار X على Y .

(٨) الجدول التالي يمثل درجات 100 طالبًا في مادتي الرياضيات (X) والفيزياء (Y) .

y \ x	40 – 49	50 – 59	60 – 69	70 – 79	80 – 89	90 – 99
90 – 99				2	4	4
80 – 89			1	4	6	5
70 – 79			5	10	8	1
60 – 69	1	4	9	5	2	
50 – 59	3	6	6	2		
40 – 49	3	5	4			

ا) أوجد معامل الارتباط بين Y و X .

ب) أوجد خط انحدار Y على X .

(٩) الجدول التالي يبين التقديرات التي حصل عليها 480 طالبًا في اختيارين مختلفين والمطلوب إيجاد معامل التوافق بين تقديرات الطلبة في المادتين.

الأول / الثاني	ممتاز	جيد	مقبول	المجموع
مقبول	10	20	100	130
جيد	30	170	40	240
ممتاز	60	30	20	110
المجموع	100	220	160	480

(١٠) الجدول التالي يبين عدد الأشخاص المتعلمين وغير المتعلمين موزعين حسب ممارستهم لعادة التدخين. والمطلوب حساب معامل الاقتران.

التدخين / التعليم	لا يدخن	يدخن	المجموع
متعلم	18	5	23
غير متعلم	12	20	32
المجموع	30	25	55

٦

المجموعات
Sets

● تعريف المجموعة ● طرق تمثيل المجموعات ● المجموعة الجزئية ● المجموعة الخالية ● المجموعة الشاملة ● الاتحاد ● التقاطع ● المجموعة المكمّلة ● الفروق ● المجموعة المنتهية والمجموعة القابلة للعدّ ● مجموعات حاصل الضرب ● فصول المجموعات ● تمارين

سوف نحتاج في دراستنا لمبادىء علم الاحتمالات بعض خصائص المجموعات، والعمليات التي تتم عليها، والتي سبق دراستها في مراحل التعليم السابقة للجامعة. وسوف نتناول ما نحتاجه منها في دراسنا فيمايلي.

(٦ ـ ١) تعريف المجموعة

المجموعة هي تجمع لأشياء معرّفة تعريفًا جيدًا. والمقصود بالتعريف الجيد هو إعطاء الصفات المشتركة والمميزة للعناصر التي تشتمل عليها المجموعة. بحيث يمكن الحكم على عنصر ما بأنه ينتمي أو لا ينتمي إلى هذه المجموعة.

مثال ذلك مجموعة الدارسين لمقرر ١٠١ إحص، ففي هذه المجموعة كل طالب يدرس مقرر ١٠١ إحص يعتبر عنصرًا في هذه المجموعة. وأي طالب لا يدرس مقرر

١٠١ إحص فهو ليس عنصرًا في هذه المجموعة. وعادة يرمز للمجموعات بالرموز A, B, C, ... وللعناصر بالرموز ..... a, b, c, . فإذا كان العنصر a ينتمي إلى المجموعة A فإنه يكتب على الصورة $a \in A$ (ويقرأ a ينتمي إلى A). وإذا كان العنصر a لا ينتمي إلى المجموعة A فإنه يكتب على الصورة $a \notin A$ (ويقرأ a لا ينتمي إلى A) فمثلاً إذا كان $A = \{1,2,4\}$ فإن $2 \in A$ و $3 \notin A$.

(٦ - ٢) طرق تمثيل المجموعات

هناك طرق كثيرة لكتابة المجموعات، نذكر منها طريقة جدولة العناصر، وطريقة الصفة المميزة للعناصر، وكذلك تمثيل المجموعة بما يسمى بشكل فن (Venn). وسوف نتناول كل طريقة على حِدَه فيما يلي:

(٦ - ٢ - ١) طريقة جدولة العناصر

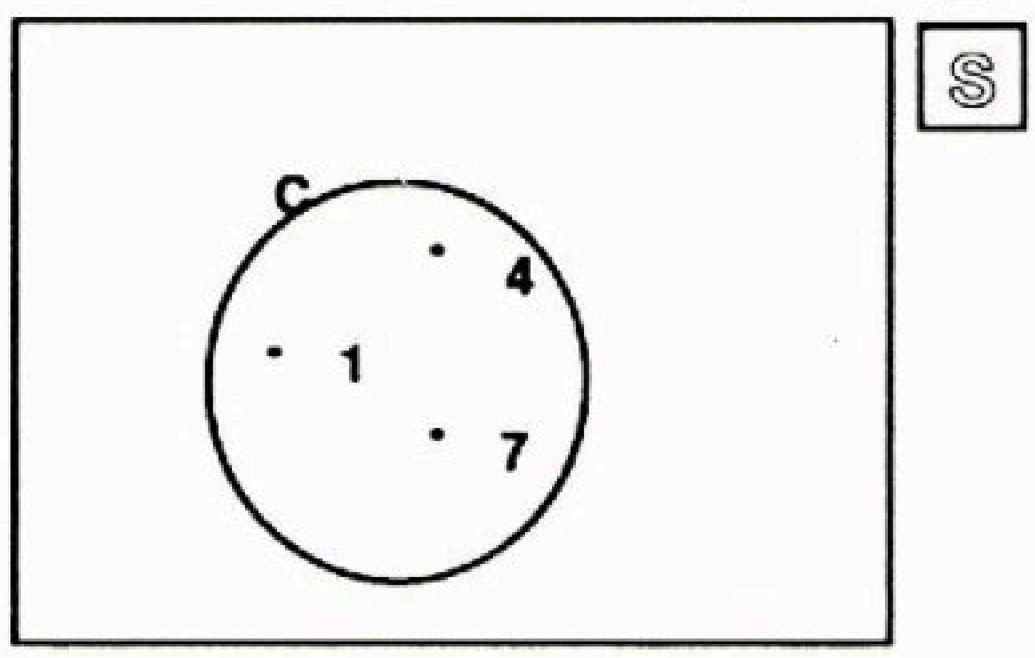
ومن هذه الطريقة يكتب اسم المجموعة ثم نكتب يساوي ثم توضع العناصر بين قوسين مثل { } وكل عنصر يفصل بعلامة فاصلة (,). مثال ذلك إذا كانت المجموعة A عبارة عن الأرقام 1, 2, 5, 6 فإنها تكتب على الصورة: $A = \{1, 2, 5, 6\}$. ويهمنا أيضًا في دراستنا للمجموعات أن نعرف عدد عناصر كل مجموعة وسوف نرمز لعدد العناصر بالرمز $n(\quad)$ حيث يكتب بين القوسين السابقين اسم المجموعة. وبذلك يكون عدد عناصر المجموعة A السابقة والذي يرمز له بالرمز n(A) يساوي 4 أي أن $n(A) = 4$.

(٦ - ٢ - ٢) طريقة الصفة المميّزة للعناصر

وفي هذه الطريقة تكتب المجموعة A على الصورة $A = \{x : h(x)\}$ حيث h(x) يمثل صفة العناصر (وتقرأ المجموعة A عبارة عن مجموعة العناصر x والتي تحقق العلاقة {h(x)} . وعلى سبيل المثال $A = \{x : x$ عدد زوجي$\}$. ويمكن كتابة هذه المجموعة A بطريقة جدولة العناصر كالتالي: $A = \{2, 4, 6, ...\}$. وفي هذا المثال أمكن كتابة المجموعة المعطاة بالصفة المميزة بطريقة جدولة العناصر، ومثال ذلك المجموعة B المعطاة بالقيمة التالية: $B = \{x : 0 \leq x \leq 3\}$.

(٦ - ٢ - ٣) أشكال فن (Venn)

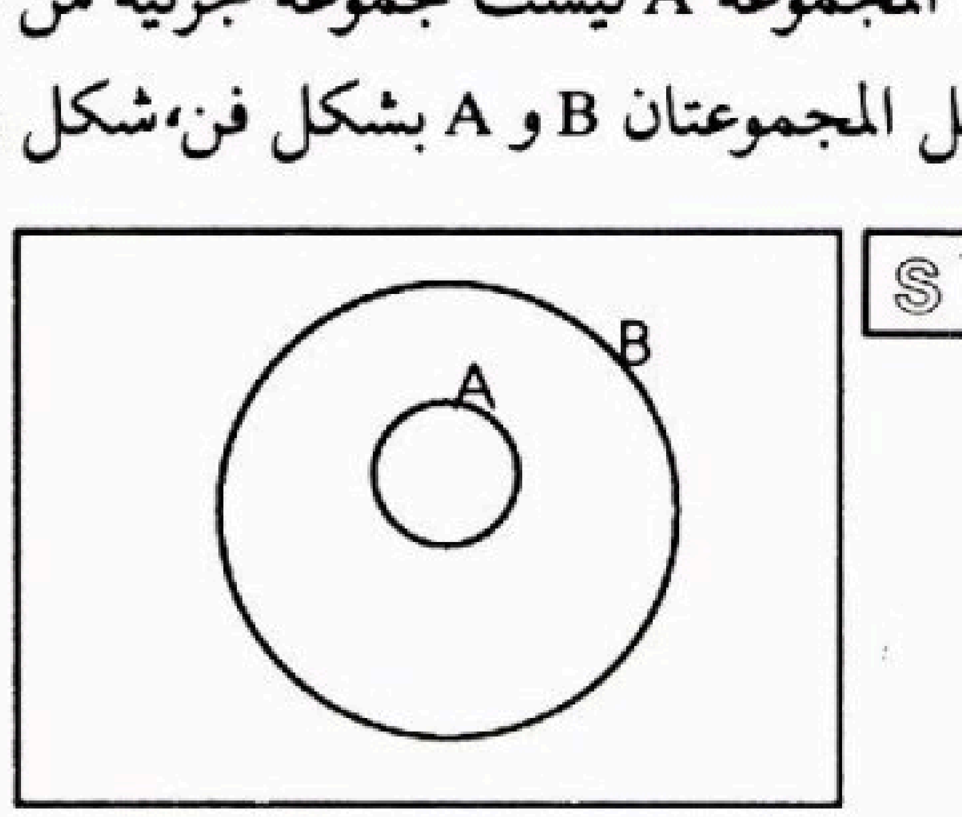
عبارة عن رسوم هندسية كالدائرة والمستطيل والمثلث. . . . الخ . ويوجد بداخلها عدة نقاط تمثل العناصر. وسوف نستخدم الدائرة لتمثيل المجموعة . وأشكال فن تساعدنا كثيراً في تفهم بعض العمليات الرياضية للمجموعات فإذا كانت المجموعة C معطاة بطريقة جدولة العناصر كالتالي $C = \{1, 4, 7\}$. فإنه يمكن تمثيلها بطريقة شكل فن كما هو موضح بالرسم،شكل (٦ - ١) التالي :



شكل (٦ - ١) يمثل المجموعة C وعناصرها داخل شكل فن

(٦ - ٣) المجموعة الجزئية (Subset)

إذا كان لدينا المجموعة A مثلاً وكانت جميع عناصرها ضمن عناصر مجموعة أخرى B مثلاً . فإنه يقال إن المجموعة A هي مجموعة جزئية من B وتكتب على الصورة $A \subset B$ [وتقرأ A مجموعة جزئية من B وباللغة الإنجليزية (A subset of B)]، أو تكتب على الصورة $B \supset A$ [وتقرأ B تحتوي على A وباللغة الإنجليزية (B contains A)] . والرمز $A \not\subset B$ يعني أن المجموعة A ليست مجموعة جزئية من B أي أن B لا تحتوي على جميع عناصر A . وتمثل المجموعتان B و A بشكل فن،شكل (٦ - ٢) التالي :



شكل (٦ - ٢) يمثل المجموعة $A \subset B$ بشكل فن

ويلاحظ أن

$$n(A) \leqslant n(B)$$

ويقال لأي مجموعة بأنها مجموعة جزئية من نفسها أي أن $A \subset A$ فإذا كان لدينا $C=\{1,3,5\}$, $A=\{1,4,7\}$, $B=\{1,2,4,7\}$ فإن $n(A)=3 \leqslant n(B)=4$, $C \not\subset B$, $A \subset B$

(٦ ـ ٤) المجموعة الخالية Empty set (ϕ)

وهي المجموعة التي لا تحتوي على أية عناصر. ويرمز لها بالرمز $\phi = \{\quad\}$ ويلاحظ أن في داخل القوسين لا يوجد أي عنصر. وكمثال للمجموعات الخالية هو مجموعة الطلّاب الذين يدرسون في كلية البنات بجامعة الملك سعود، ومجموعة الطلّاب بجامعة الملك سعود وأعمارهم أقل من سبع سنوات ويلاحظ أن $n(\phi)=0$.

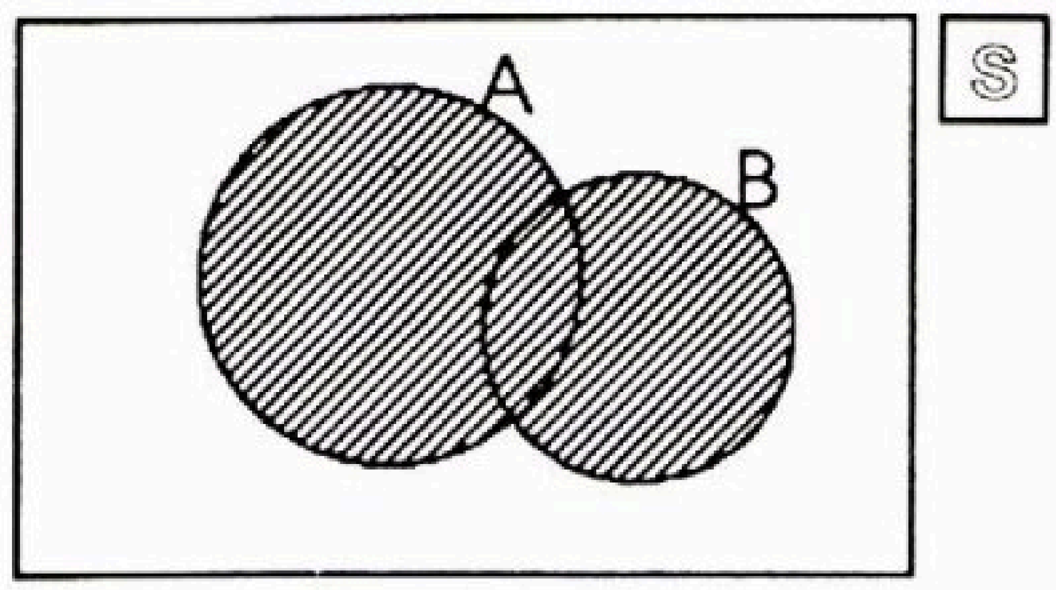
(٦ ـ ٥) المجموعة الشاملة (Universal set)

يلاحظ أن أي مجموعة من المجموعات يكون لها مجموعة أعم وأشمل منها، وهذه المجمـوعة تسمى بالمجموعة الشاملة ويرمز لها بالرمز S فمثلاً مجموعة طلّاب جامعة الملك سعـود مجمـوعـة جزئية من مجموعة طلّاب جامعات المملكة، وكذلك مجموعة طلاب جامعات المملكة تعتبر مجموعة جزئية من طلّاب جامعات الدول الإسلامية، وهكذا. . . . ، وخلاصة القول فالمجموعة الشاملة هي المجموعة التي تحتوي على كل العناصر.

وسوف تمثل المجموعة الشاملة بشكل فن بمستطيل، ترسم داخله دوائر تمثل المجموعة الجزئية من هذه المجموعة الشاملة.

(٦ ـ ٦) الاتحـاد (Union)

يعرف الاتحاد لمجموعتين B و A مثلاً بأنه عبارة عن المجموعة C التي تمثل مجموعة العناصر الموجودة في A أو B أو كليهما معًا. ويرمز للاتحاد بالرمز $C=A \cup B$ (ويقرأ A اتحاد B : A union B) وتمثل بشكل فن،شكل (٦ ـ ٣) التالي:

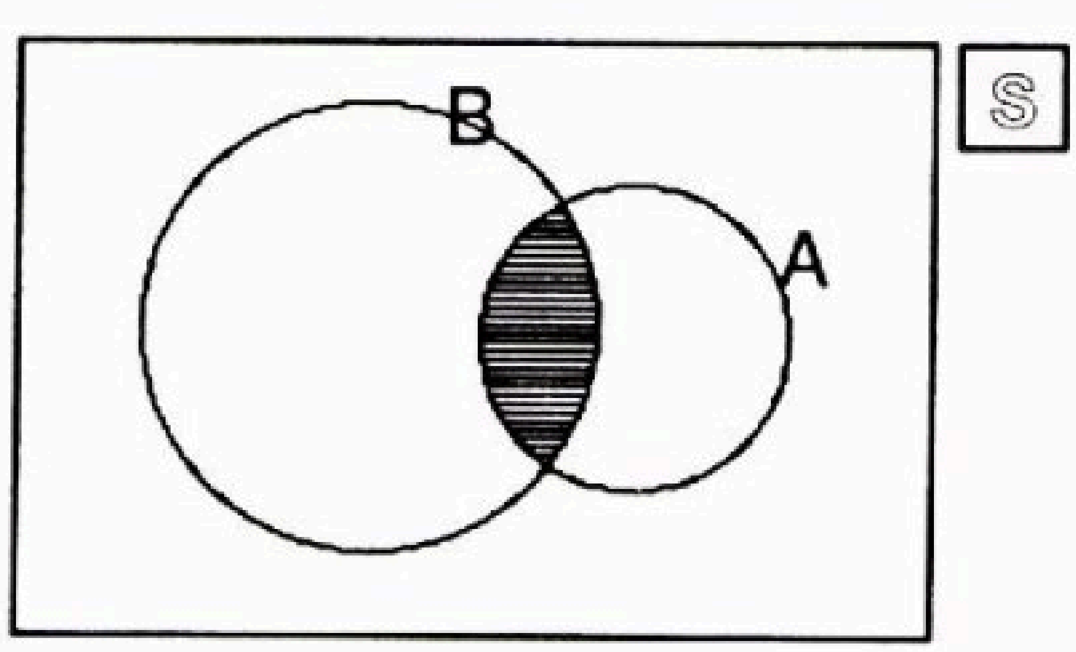


شكل (٦ ـ ٣) يمثل A∪B شكل فن

والجزء المظلّل في الشكل السابق (٦ ـ ٣) هو عبارة عن A∪B .

(٦ ـ ٧) التقاطع (Intersection)

يُعرّف تقاطع مجموعتين B و A مثلاً بأنه مجموعة العناصر C التي تتكون من عناصر B و A المتشابهة معًا ويكتب على الصورة C=A∩B ويقرأ A تقاطع B [A inter-section B] ويوضح بشكل فن،شكل (٦ ـ ٤) التالي :



شكل (٦ ـ ٤) يمثل A∩B بشكل فن

والجزء المظلل في الشكل السابق (٦ ـ ٤) هو عبارة عن A∩B .

مثال (٦ ـ ١)

إذا كان :

$$A = \{1,2,4,7\} \ , \ B = \{2,5,7\}$$
$$C = \{3,10\}$$

أوجد قيم كل من : $A \cup B \ , \ A \cap B \ , \ B \cap C$

ثم أوجد عدد عناصر كل من المجموعات الجديدة المكونة من الاتحاد والتقاطع .

الحـــل

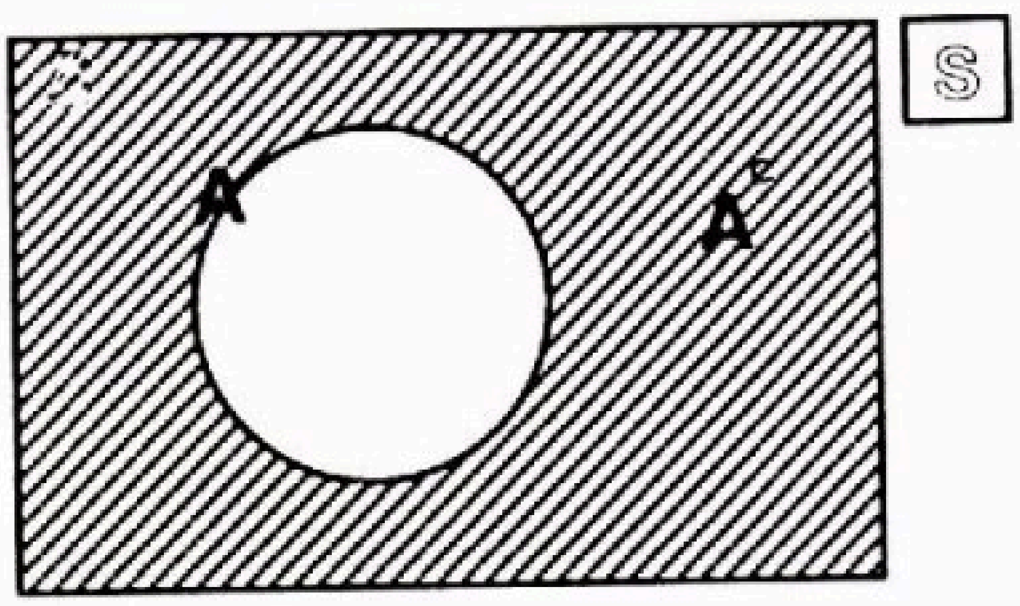
$A \cup B = \{1,2,4,5,7\}$ $\qquad n(A \cup B) = 5$

$A \cap B = \{2,7\}$ $\qquad n(A \cap B) = 2$

$A \cap C = \phi$ $\qquad n(B \cap C) = 0$

(٦ ـ ٨) المجموعة المكمّلة (Complement set)

تعرف المجموعة المكملة لأي مجموعة A مثلاً بأنها مجموعة العناصر الموجودة في المجموعة الشاملة S وليست في المجموعة A نفسها ويرمز لها بالرمز A^c أو $\bar{A}$ [وتقرأ مكملة A : A complement] وتوضح في شكل فن، شكل (٦ ـ ٥) التالي بالجزء المظلّل .



شكل (٦ ـ ٥) يمثل $\bar{A}$

ونلاحظ أن $\quad A \cap A^c = \phi \ , \ A \cup A^c = S$

وأن عدد عناصر A^c هو $n(A^c)$ ويعطى بالتالي :

$$n(A^c) = n(S) - n(A)$$

مثال (٦ ـ ٢)

إذا كانت $S = \{1,3,5,6,7\}\ ,\ A = \{3,6\}$

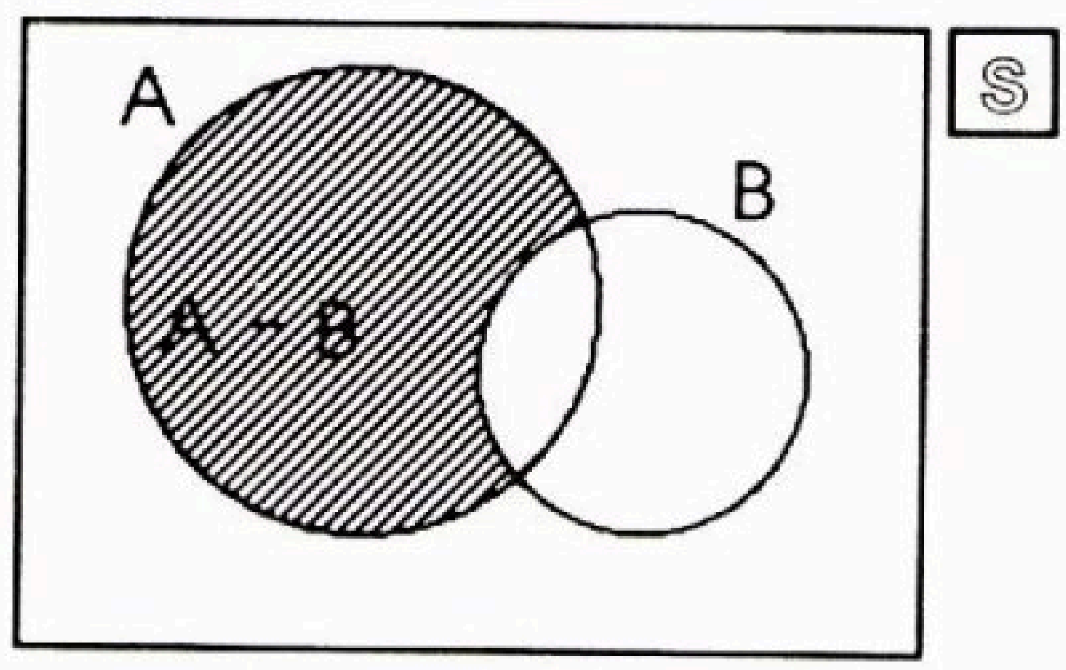
فأوجد A^c **و** $n(A^c)$.

الحـــل

$$A^c = \{1,5,7\}\ ,\ n(A^c) = 3$$

(٦ ـ ٩) الفـــروق (Differences)

إذا كان لدينا مجموعتان B و A فإن A − B يسمى بالفرق وهو عبارة عن المجموعة التي توجد عناصرها في A وليست في B كما هو موضح بالجزء المظلل في الشكل (٦ ـ ٦) التالي :



شكل (٦ ـ ٦) يمثل (A − B) بشكل فن

ويمكن بسهولة إثبات أن المجموعة A − B هي المجموعة $A \cap B^c$ وكذلك المجموعة B − A هي المجموعة $B \cap A^c$.

مثال (٦ ـ ٣)

إذا كان $S = \{1,4,5,6,7,8\}\ ,\ A = \{1,4,7\}\ ,\ B = \{5,6,7\}$

أوجد مجموعة الفرق A − B بطريقتين.

الحـــل

أولاً

$$B^c = S - B$$

$$= \{1,4,8\}$$

$$A \cap B^c = \{1,4\}$$

ثانيًا:

من تعريف الفرق وهو العناصر الموجودة في A وليست في B ينتج أن مجموعة الفرق تكون كالتالي:

$$A - B = \{1,4\}$$

وهي النتيجة السابقة نفسها.

أي أن:

$$A - B = A \cap B^c$$

مثال (٦ ـ ٤)

إذا كان

$$S = \{1,2,3,4,5,6\} \ , \ A = \{1,3,5\} \ , \ B = \{2,4,6\} \ , \ C = \{1,2,4\}$$

أوجد

$$A^c, B^c, C^c, A\cup B, B\cup C, A\cup B\cup C, A^c\cap B^c, (A\cup B)^c, A^c\cup B^c$$

$$(A\cap B)^c, A\cap(B\cup C), (A\cap B) \cup (A\cap C), A\cup(B\cap C), (A\cup B) \cap (A\cup C)$$

الحـــل

$$A^c = \{2,4,6\} \ , \ B^c = \{1,3,5\}, C^c = \{3,5,6\}$$

$$A\cup B = \{1,2,3,4,5,6\}$$

$$B\cup C = \{1,2,4,6\}$$

$$A\cup B\cup C = \{1,3,5\}\cup\{1,2,4,6\}$$

$$= \{ 1,2,3,4,5,6\}$$

$$A^c\cap B^c = \{2,4,5\} \cap \{1,3,5\} = \phi$$

$$(A\cup B)^c = \{1,2,3,4,5,6\}^c = \phi$$

$$\therefore A^c\cap B^c = (A\cup B)^c \qquad \text{.......................} \quad (٦ - ١)$$

$$A\cap B = \{ \qquad \} = \phi$$

$$\because A\cap C = \{1\}$$

$$A^c\cup B^c = \{2,4,6\} \cup \{1,3,5\}$$

$$= \{1,2,3,4,5,6\}$$

$$(A\cap B)^c = \{\quad\}^c = \phi^c = \{1,2,3,4,5,6\}$$

$$\therefore\ A^c\cup B^c = (A\cap B)^c \quad \text{.........} \quad (٦ - ٢)$$

ملاحظة

كل من العلاقة (٦ ـ ١)، (٦ ـ ٢) تسمى بقانون دي مورجان (Demorgan)

$$A\cap(B\cup C) = \{1,3,5\} \cap \{1,2,4,6\}$$
$$= \{1\}$$
$$(A\cap B) \cup (A\cap C) = \{\quad\} \cup \{1\} = \{1\}$$
$$\therefore\ A\cap (B\cup C) = (A\cap B) \cup (A\cap C) \quad \text{.........} \quad (٦ - ٣)$$

ملاحظة

العلاقة (٦ ـ ٣) تسمى بقانون توزيع التقاطع على الاتحاد.

$$A\cup (B\cap C) = \{1,3,5\} \cup \{2,4\}$$
$$= \{1,2,3,4,5\}$$
$$(A\cup B) \cap (A \cup C) = \{1,2,3,4,5,6\} \cap \{1,2,3,4,5\}$$
$$= \{1,2,3,4,5\}$$
$$\therefore\ A\cup (B\cap C) = (A\cup B) \cap (A\cup C) \quad \text{.........} \quad (٦ - ٤)$$

ملاحظة

العلاقة (٦ ـ ٤) تسمى بقانون توزيع الاتحاد على التقاطع ويمكن تلخيص القوانين السابقة كمايلي:

(١) قانون دي مورجان

$$(A\cup B)^c = A^c\cap B^c$$
$$(A\cap B)^c = A^c\cup B^c$$

(٢) قانون التبادل

$$A \cap B = B \cap A$$

$$A \cup B = B \cup A$$

(٣) قانون التجميع

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

(٤) قانون التوزيع

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (B \cap C)$$

(٥) قانون الاتمام (complement)

$$A^c \cup A = S \ , \ A \cup S = S$$

$$A^c \cap A = \phi \ , \ A \cap S = A$$

(٦ - ١٠) المجموعة المنتهية والمجموعة القابلة للعدّ

المجموعات المنتهية والمجموعات القابلة للعدّ مهمة جدًّا في دراسة علم الاحتمال. ويقال إن المجموعة منتهية إذا كانت مجموعة خالية أو تحتوي على عدد من العناصر مقداره n عنصرًا. حيث إن n عدد صحيح موجب وأقل من ما لا نهاية $(0 \leqslant n < \infty)$. وجميع المجموعات التي وردت في الأمثلة السابقة (٦ - ١)، (٦ - ٢)، (٦ - ٣) هي عبارة عن مجموعات منتهية. وتسمى المجموعة لا نهائية إذا كانت غير منتهية. ويقال إن المجموعة قابلة للعدّ إذا كانت هذه المجموعة منتهية أو مجموعة لا نهائية ولكن يمكن وضع عناصرها على صورة متتابعة مثل مجموعة الأعداد الفردية الموجبة أو مجموعة الأعداد الطبيعية. . . الخ. وتوصف المجموعة اللانهائية والتي لا يمكن وضع عناصرها في صورة متتابعة بأنها مجموعة لا نهائية غير قابلة للعدّ مثل مجموعة A للأعداد الحقيقية في الفترة (0,1) وتكتب $A = \{x : 0 < x < 1\}$.

مثال (٦ ـ ٤)

أي المجموعات التالية منتهية أو قابلة للعدّ أو غير قابلة للعدّ.

A = { شهور العام الهجري }

B = { الأعداد الحقيقية R }

C = { الأعداد النسبية }

الحـــل

المجمـوعـة A هي مجمـوعة منتهية لأن الشهور عددها 12 شهرًا والمجموعة B لا نهائية وغير قابلة للعدّ لأنه لا يمكن وضع عناصرها على صورة متتابعة.

المجموعة C لا نهائية ولكن قابلة للعدّ لأنه يمكن وضع عناصرها في صورة متتابعة.

(٦ ـ ١١) مجموعات حاصل الضـرب

إذا كان لدينا مجموعتان B و A فإن المجموعة التي تتكون من حاصل ضرب B و A ويرمز لها بالرمز $A \times B$ وهي عبارة عن مجموعة جميع الأزواج المرتبة من a,b حيث $a \in A\ ,\ b \in B$ ويعبر عنها رياضيًا كالتالي:

$$A \times B = \{(a,b) : a \in A\ ,\ b \in B\}$$

وإذا كان حاصل الضرب عبارة عن المجموعة في نفسها مثل $A \times A$ فإنه يرمز لها بـ A^2 حيث

$$A \times A = A^2 = \{(a,b) : a \in A\ ,\ b \in A\}$$

مثال (٦ ـ ٥)

إذا كانت: $A = \{H,T\}\ ,\ B = \{1,2,4\}$

أوجد $A \times B\ ,\ B \times A\ ,\ A \times A\ ,\ B \times B$

وعدد عناصر كل منهم.

الحـــل

$$A \times B = \{(H,1)\ ,(H,2)\ ,(H,4)\ ,(T,1)\ ,(T,2)\ ,(T,4)\}$$

$$B \times A = \{(1,H)\ ,(1,T)\ ,(2,H)\ ,(2,T)\ ,(4,H)\ ,(4,T)\}$$

$$A \times A = \{(H,H), (H,T), (T,H), (T,T)\}$$
$$B \times B = \{(1,1), (1,2), (1,4), (2,1), (2,2), (2,4), (4,1), (4,2), (4,4)\}$$
$$n(A \times B) = 6 \quad , \quad n(B \times A) = 6$$
$$n(A \times A) = 4 \quad , \quad n(B \times B) = 9$$

(٦ - ١٢) فصول المجموعات (Classes of sets)

يعـرف فصـل المجمـوعـات الجزئية $\xi(A)$ بأنه المجموعة التي تكون عناصرها مجموعات جزئية من المجموعة A .

فمثلاً إذا كان لدينا المجموعة $A = \{a,b,c\}$

فإن فصل المجموعات الجزئية $\xi(A)$ إذا احتوى على جميع المجموعات الجزئية من A فإنه يسمى **مجموعة القوى لـ** A ويرمز له بـ (P_A) ويكون

$$P_A = \{\phi , \{a\} , \{b\} , \{c\} , \{a,b\} , \{a,c\} , \{b,c\} , \{a,b,c\}\}$$

نلاحظ أن عدد عناصر المجموعة A هو $n(A) = 3$.

وأن عدد عناصر P_A هو $n(P_A) = 8$ أي يساوي 2^3 .

وعمومًا عدد عناصر P_A هو $2^{n(A)}$.

(٦ - ١٢ - ١) تجزئة المجموعة

إذا كان لدينـا مجمـوعـة A مثـلاً وقسمت إلى مجمـوعـات جزئية غير خالية $A_1, A_2, \ldots, A_n$ بحيث كان :

$$(i)\ A_1 \cup A_2 \cup \ \ldots\ldots \cup A_n = A$$
$$(ii)\ A_i \cap A_j = \phi \quad , \quad i , j = 1,2, \ldots ,n \quad , \quad i \neq j$$

فإن المجموعات $A_1, A_2, \ldots , A_n$ تسمى **تجزيء A** ، أي أن تجزيء A هو فصل من المجموعات الجزئية غير الخالية للمجموعة A بحيث ينتمي كل $a \in A$ إلى مجموعة جزئية واحدة فقط وتسمى **المجموعة الجزئية في التجزيء بالخلايا**.

مثال (٦ - ٦)

إذا كانت المجموعة A كالتالي :

$$A = \{2,4,5,6,8,9\}$$

أي من الفصول التالية يعتبر تجزيئًا للمجموعة A وأي منها ليس تجزيئًا؟ مع ذكر السبب؟

$$\text{i)}\ \{\{2\}, \{4\}, \{6,8,9\}\}$$
$$\text{ii)}\ \{\{2\}, \{4,5\}, \{5,6,8,9\}\}$$
$$\text{iii)}\ \{\{2,4\}, \{5\}, \{6,8,9\}\}$$

نجد أن (i) ليست تجزيئًا للمجموعة A لأن العنصر 5 للمجموعة A لا ينتمي إلى أية خلية.

وكذلك (ii) ليست تجزيئًا للمجموعة لأن العنصر $5\in A$ موجود في خليتين.

ونلاحظ (iii) هو عبارة عن تجزيء للمجموعة A لأن كل عنصر في A ينتمي إلى خلية واحدة فقط.

(٦ - ١٢ - ٢) تعريف (جبر سجما) σ - Algebra

يسمى فصل غير خالٍ ξ من المجموعات الجزئية للمجموعة الشاملة بجبر المجموعات (جبر المجموعات σ) إذا تحقق الشرطان التاليان :

١ - المجموعة المكملة لأي مجموعة في ξ تنتمي إلى ξ أيضًا.

٢ - اتحاد أي عدد غير منتهٍ وقابل للعدّ من مجموعات ξ ينتمي إلى ξ أيضًا.

أي أن الفصل ξ يحقق شرط الانغلاق لكل من المتممة والاتحاد ويحتوي على Φ ، S .

(٦ - ١٣) تمـــارين

(١) إذا كانت $A = \{x : 4x = 16\}$

أي من العلاقات التالية صحيحة مع ذكر السبب

$$\text{i)}\ A = 4 \quad , \quad \text{ii)}\ A = \{4\}$$

(٢) أي من المجموعات التالية مجموعة خالية مع ذكر السبب لكل مجموعة

i) $A = \{x : x + 2 = 2\}$

ii) $B = \{x : x \neq x\}$

iii) $C = \{x : x^3 = 27 , x = 5\}$

(٣) أي من المجموعات التالية متساوية مع ذكر السبب

$\{a,b,c\} \quad , \quad \{b,a,c\} \quad , \quad \{a,b,d\}$

(٤) أثبت أن المجموعة $D = \{1,3,5,6\}$

ليست مجموعة جزئية من $B = \{x :$ عدد فردي $x\}$

(٥) إذا كانت $S = \{a,b,c,d,e,f,g\} \ , \ A = \{a,b\} \ ,$

$B = \{e,f,g\} \ , \qquad C = \{a,g\}$

أوجد التالي: $A^c, B^c, C^c, A \cup B, A \cap B^c, B \cap A^c,$

$A \cap B \cap C, A \cup B \cup C, A \cap C, A \cap B$

(٦) إذا كان $A = \{1,2,3,4,5,6\}$

أوجد: $A \times A$

(٧) إذا كان: $A = \{\{2,4\} \ , \ B = \{6,10\} \ , \ C = \{8,12\}$

أوجد $A \times (B \cup C) \ , \ (A \cup B) \ , \ (B \cup C)$

$(A \times B) \cup (B \times C) \quad , \ (A \times B) \cap (B \times C)$

(٨) إذا كان: $A = \{1,2,3,4,5,6\}$

i) $A_1 = \{1,3,5\} \ , \ A_2 = \{2\} \quad , \ A_3 = \{4,8\}$

ii) $B_1 = \{1,4\} \quad , \ B_2 = \{3,5\} \ , \ B_3 = \{2,6\}$

أي من الفصول التالية يكون تجزيئًا للمجموعة A ؟

$\{A_1, A_2, A_3\} \quad , \quad \{B_1, B_2, B_3\}$

(٩) إذا كان لدينا مجموعة منتهية A عدد عناصرها n .

فأثبت أن عدد عناصر الفصل $\xi(A)$ هو $n\{\xi(A)\} = 2^n$.

٧

طرق العد

Counting

● القاعدة الأساسية لطرق العدّ (قاعدة الضرب ، قاعدة الجمع) ● التباديل ● تطبيق على التباديل ● التوافيق ● تمارين

لقد سبق لنا دراسة المجموعات وعرفنا كيف يمكن إيجاد عدد عناصرها . وفي معظم الأحوال يكون عدد العناصر كبيراً جدًّا ويتطلب جهدًا كبيراً في كتابته وتكون عرضة للخطأ أثناء كتابته وعدّه . لذلك دعت الحاجة إلى إيجاد طرق رياضية أخرى تساعدنا في إيجاد عدد عناصر المجموعة دون الحاجة إلى كتابة المجموعة بطريقة جدولة العناصر أو غيرها . وهذه الطرق تساعدنا أيضًا في إيجاد عدد الحالات الممكنة لأي تجربة وهي مفيدة في علم الاحتمالات كما سوف يتضح فيما بعد .

(٧ ـ ١) القاعدة الأساسية لطرق العَدّ (قاعدة الضرب وقاعدة الجمع)

إذا كان لدينا عملية تتم على مرحلتين وكانت تتم في المرحلة الأولى بعدد قدره n_1 طريقة وفي المرحلة الثانية تتم بعدد قدره n_2 طريقة فإن عدد الطرق الممكنة لإتمام هذه العملية في المرحلتين معًا هو $n_1 \times n_2$ طريقة ونوضح ذلك بالأمثلة التالية .

مثال (٧ - ١)

بكم طريقة يمكن أن يختار طالب مقررين الأول للإحصاء والثاني للرياضيات إذا كان يوجد بقسم الإحصاء 5 مقررات تناسب هذا الطالب وبقسم الرياضيات 3 مقررات تناسبه أيضًا.

الحــل

عدد طرق اختيار الطالب لمقرر من الإحصاء = 5 طرق.
عدد طرق اختيار الطالب لمقرر من الرياضيات = 3 طرق.
عدد طرق اختيار الطالب لمقررين الأول من الإحصاء والثاني من الرياضيات
$= 5 \times 3 = 15$ طريقة.

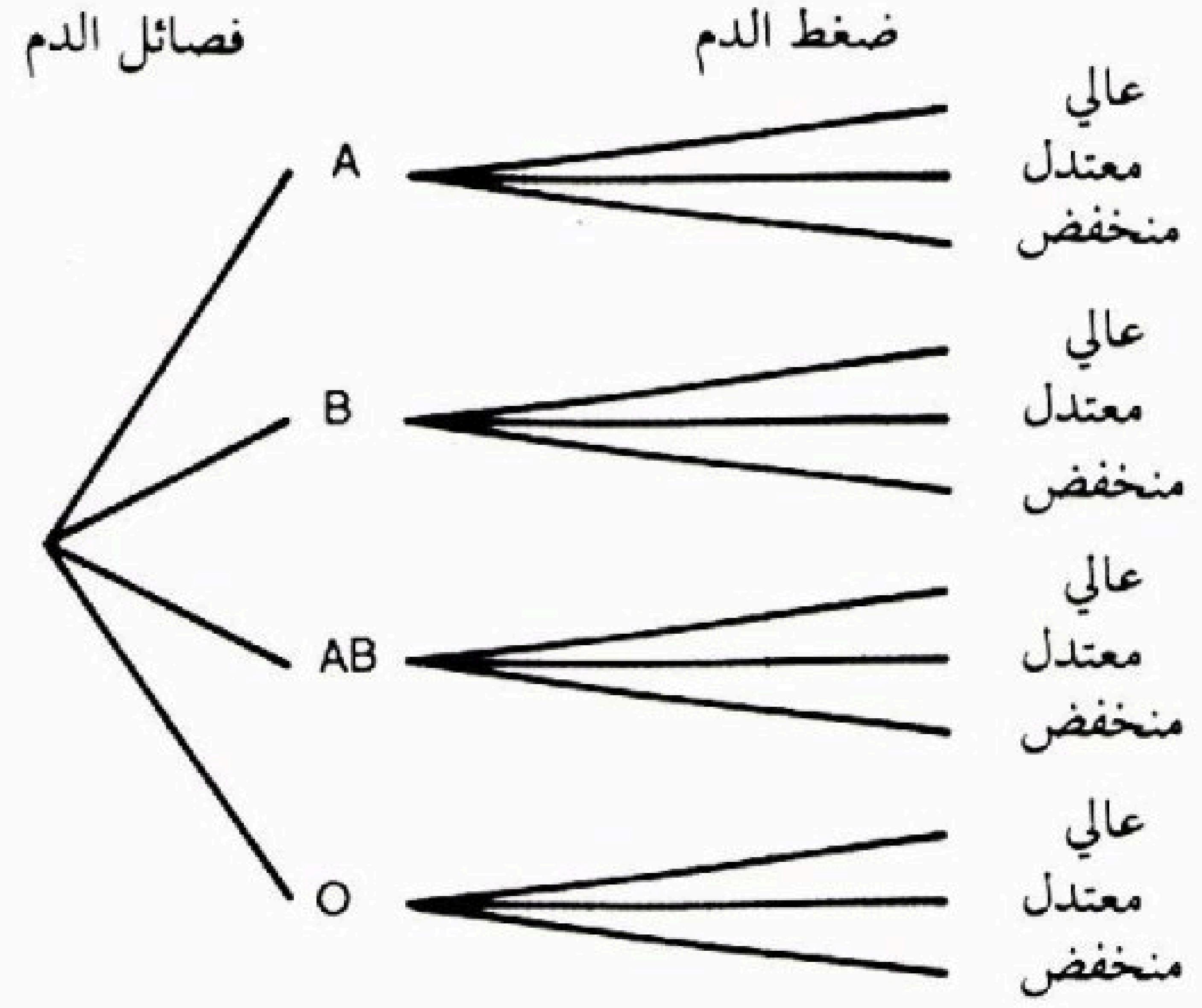
مثال (٧ - ٢)

عند القيام بدراسة طبيّة صنّفنا المريض طبقًا لنوع فصيلة الدم وهي A, B, AB, O وكذلك بالنسبة لضغط الدم وهو مرتفع، معتدل، منخفض فبكم طريقة يتم تصنيف المريض من حيث فصيلة الدم وضغط الدم معًا؟

الحــل

عدد طرق فصائل الدم = 4
عدد طرق أنواع ضغط الدم = 3
عدد طرق تصنيف المريض بالنسبة لفصيلة الدم وضغط الدم معًا
$= 4 \times 3 = 12$.

ويمكن توضيح الحل السابق باستخدام الشجرة البيانية التالية في شكل (٧ - ١):



شكل (٧ ـ ١) يُمثّل الشجرة البيانية لفصائل الدم وضغط الدم

فيكون عدد الطرق كالتالي:

(منخفض ,B) و(معتدل ,B) و(عالي ,B) و(منخفض ,A) و(معتدل ,A) و(عالي ,A) و(منخفض ,O) و(معتدل ,O) و(عالي ,O) و(منخفض ,AB) و(معتدل ,AB) و(عالي ,AB).

ويكون عدد الطرق = 12 .

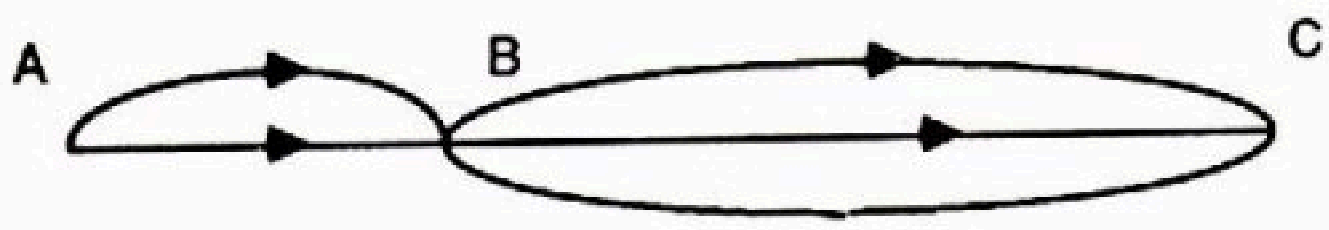
وهي النتيجة الأولى نفسها = 4×3 طريقة.

مثال (٧ ـ ٣)

إذا كان لدينا ثلاث مدن A, B, C وكان هناك طريقان من A إلى B وثلاث طرق من B إلى C فإذا قام شخص من A إلى C مارًا بالمدينة B فبكم طريقة يمكن أن يصل هذا الشخص من المدينة A إلى المدينة C .

الحــل

نوضح الطرق بين المدن الثلاث A, B, C بالرسم، شكل (٧ ـ ٢) التالي:



شكل (٧ ـ ٢) يوضح الطرق المختلفة بين المدن الثلاث

الشخص يمكن أن يصل من A إلى B بعدد طرق قدره طريقان ويصل أيضًا من B إلى C بعدد طرق قدره 3 طرق.

عدد الطرق للشخص من A إلى C = 2 x 3 = 6 طرق.

مثال (٧ ـ ٤)

ذهب خالد لشراء سيارة من أحد المعارض فوجد فيه أربعة موديلات مختلفة. لها ثلاثة أنواع مختلفة من المحركات ولها ألوان مختلفة عددها 12 لونًا. كم من الخيارات المختلفة (من حيث الموديل واللون والمحرك) المتاحة لاختيار سيارة من المعرض؟

الحـــل

عدد طرق اختيار الموديل = 4 طرق.

عدد طرق اختيار المحرك = 3 طرق.

عدد طرق اختيار اللون = 12 طريقة.

عدد طرق اختيار خالد لواحدة من السيارات بالمعرض = 12x3x4 = 144 طريقة.

ويمكن تعميم عدد طرق الاختيار في r من المراحل كالتالي:

إذا كان لدينا الاختيار في المرحلة الأولى بعدد قدره n_1 طريقة وفي المرحلة الثانية n_2 طريقـة . . . . . والمـرحلة الأخـيرة n_r طريقة فيكون عدد الطرق المختلفة الكلية

$= n_1 \times n_2 \times \ldots \times n_r$ طريقة وتسمى هذه بالقاعدة الأساسية للعد (أو قاعدة الضرب) كما توجد قاعدة أخرى تسمى بقاعدة الجمع كمايلي:

إذا كانت العملية A تتم بطريقة واحدة من n_1 طريقة والعملية B تتم بطريقة واحدة من n_2 طريقة وكانت العمليتان متنافيتان فإن عدد طرق اتمام عملية منها أو الأخرى هو حاصل الجمع $n_1 + n_2$.

مثال (٧ ـ ٥)

بكم طريقة يمكن أن يختار طالب مقررًا واحدًا من الإحصاء أو الرياضيات. إذا كان يوجد بقسم الإحصاء 3 مقررات تناسب هذا الطالب وبقسم الرياضيات 4 مقررات تناسبه أيضًا.

الحــل

عدد طرق اختبار الطالب لمقرر الإحصاء = 3 طرق.
عدد طرق اختبار الطالب لمقرر من الرياضيات = 4 طرق.
عدد طرق اختبار الطالب لمقرر واحد من الإحصاء أو من الرياضيات= $4 + 3 = 7$ طرق.

ويمكن تعميم القاعدة السابقة في r من العمليات المتنافية كالتالي:

إذا كان الاختيار في العملية الأولى بعدد طرق قدره n_1 طريقة وفي العملية الثانية n_2 طريقة. . . . . . والعملية الأخيرة بعد n_r طريقة فيكون عدد الطرق الكلية = $n_1 + n_2 + \ldots + n_r$ طريقة. وتسمى هذه بالقاعدة الأساسية للعدّ أيضًا (أو قاعدة الجمع).

(٧ ـ ٢) التباديل (Permutation)

هي ترتيبة لعدة أشياء مختلفة بأخذها كلها أو بعضها في كل مرة مع مراعاة

الترتيب. وعدد تباديل n من الأشياء مأخوذة r في كل مرة يساوي عدد الترتيبات التي يمكن تكوينها من n من الأشياء بحيث يحتوي كل ترتيب على r من هذه الأشياء. وقبل كتابة قانون التباديل نعتبر المثال التالي:

مثال (٧ - ٦)

إذا كان لدينا الحروف الثلاثة a, b, c أوجد عدد الطرق الممكنة لتكوين حرفين من هذه الحروف بحيث لا يتكرر أي حرف أي مرة.

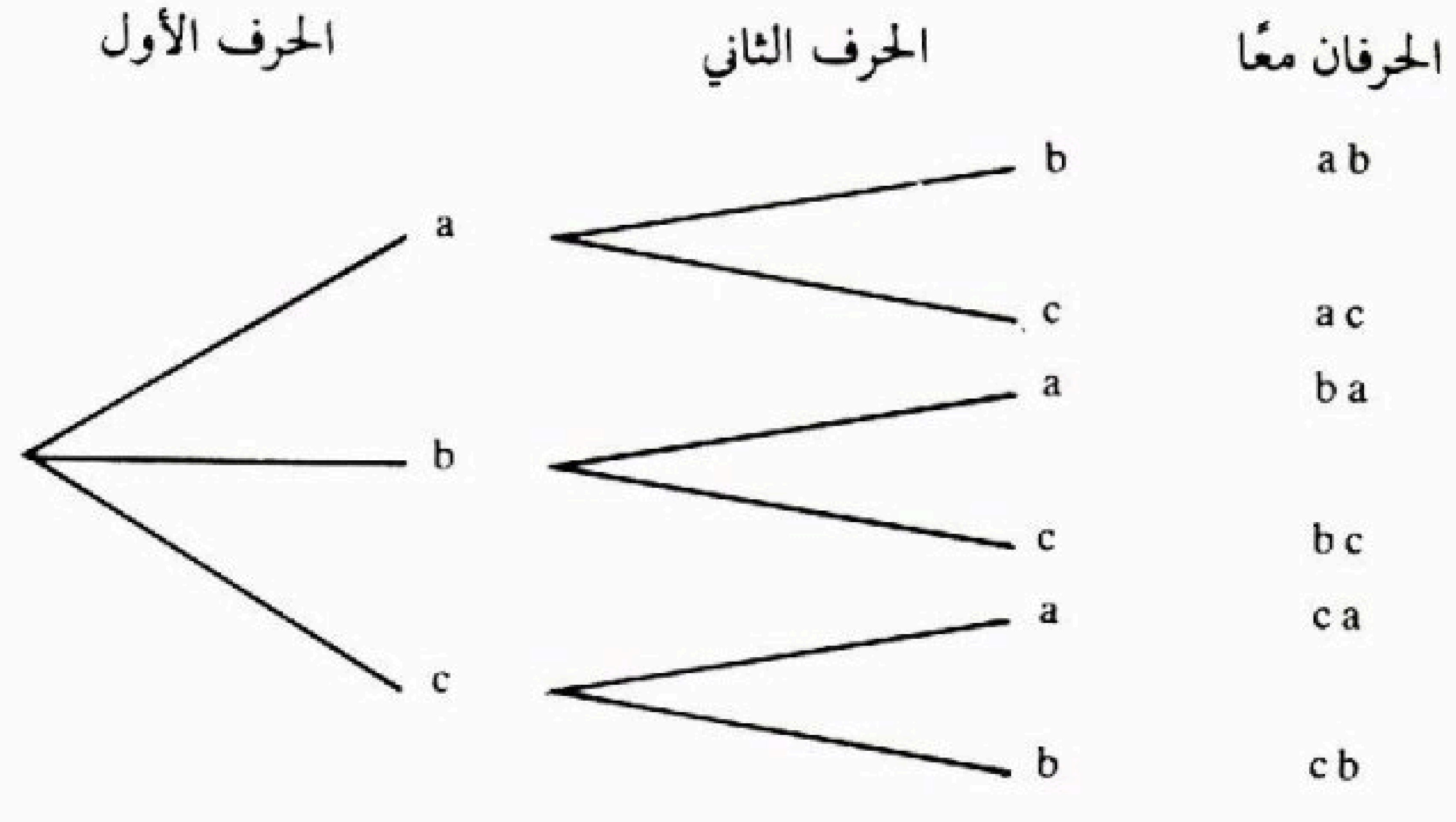
الحـــل

يكون عدد التراتيب أو التباديل هو:

ab, ac, bc, cb, ca, ba

عدد التراتيب (التباديل) = 6 ترتيبة.

ويمكن توضيح ذلك بالشجرة البيانية التالية شكل (٧ - ٣):



شكل (٧ - ٣) يمثل الشجرة للتباديل المختلفة للحروف a, b, c

عدد طرق ترتيب حرفين من ثلاث حروف = 6

وعلى حسب القاعدة الأساسية لطرق العد يمكن اختيار الحرف الأول بطرق عددها 3 والحرف الثاني بطرق عددها 2 فيكون عدد الطرق الممكنة = 3 × 2 = 6 .

وبوجه عام عدد تباديل n من الأشياء مأخوذة (r) في كل مرة سوف يرمز له بالرمز ^nP_r ويعطى بالعلاقة التالية :

$$^nP_r = n(n-1)(n-2) \ldots\ldots (n-r+1)$$

الإثبات

يمكن اختبار العنصر الأول في الترتيبة بعدد n طريقة .

ويمكن اختيار العنصر الثاني في الترتيبة بعدد (n - 1) طريقة .

ويمكن اختيار العنصر الرائي في الترتيبة بعدد (n - r + 1) طريقة .

وباستخدام القاعدة الأساسية للعدّ فإن عدد الطرق الكلية يساوي :

$$n(n-1)(n-2) \ldots\ldots\ldots (n-r+1)$$

أي أن :

$$^nP_r = n(n-1) \ldots\ldots (n-r+1) \qquad \ldots\ldots\ldots\ldots\ldots \quad (٧ - ١)$$

وعندما r = n فإن المعادلة (٧ ـ ١) تصبح :

$$^nP_n = n(n-1)(n-2) \ldots\ldots 3 \times 2 \times 1 \qquad \ldots\ldots\ldots\ldots\ldots \quad (٧ - ٢)$$

ويرمز للطرف الأيمن للمعادلة (٧ ـ ٢) بالرمز ! n (ويقرأ مضروب n) وتصبح المعادلة (٧ ـ ٢) كالتالي :

$$^nP_n = n! \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (٧ - ٣)$$

أي أن :

$$n! = n(n-1)(n-2) \ldots\ldots 3 \times 2 \times 1$$

وتسمى ^nP_n عدد الترتيبات الممكنة لجميع الأشياء .

مثال (٧ ـ ٧)

أوجد قيم ! 5 و $^5P_2$ و $^6P_3$

الحـــل

$$^6P_3 = 6 \times 5 \times 4 = 120$$

$$^5P_2 = 5 \times 4 = 20$$

$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

ويمكن وضع العلاقة (٧ ـ ١) على الصورة التالية :

$$^{n}P_{r} = \frac{n(n-1)(n-2)\ .....\ (n-r+1)(n-r)!}{(n-r)!}$$

أي أن :

$$^{n}P_{r} = \frac{n!}{(n-r)!} \qquad \text{.......................} \qquad (٧ - ٤)$$

مثال (٧ ـ ٨)

أوجد قيمة $0!$ (مضروب الصفر)

الحــل

بوضع $r = n$ في العلاقة (٧ ـ ٤) نحصل على :

$$^{n}P_{n} = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

ولكن $^{n}P_{n} = n!$

$$n! = \frac{n!}{0!}$$

$$0! = \frac{n!}{n!} = 1$$

أي أن $0!$ يساوي الواحد الصحيح .

مثال (٧ ـ ٩)

بكم طريقة يمكن أن يجلس ثلاثة أشخاص على ثلاثة مقاعد في صف واحد؟

الحــل

المقعد الأول يمكن أن يملأ بثلاثة أشخاص والثاني يمكن أن يملأ بشخصين والثالث يملأ بشخص واحد .

عدد طرق جلوس الأشخاص الثلاثة $= 3 \times 2 \times 1 = 6$ طرق .

ويمكن الحصول على النتيجة نفسها باستخدام ${}^{n}P_{n}$ حيث $n = 3$ فيكون عدد طرق جلوس الأشخاص الثلاثة $= {}^{3}P_{3}$

$= 3!$

$= 3 \times 2 \times 1 = 6$ طرق.

(٧ ـ ٣) تطبيق على التباديل

هنـاك تطبيقـات كثـيرة للتبـاديل نذكـر بعضًـا منهـا لأهميتـه في دراسـة علم الاحتمالات. وهو إيجاد عدة طرق لاختيار شيء ما مرة أو أكثر من مرة من مجموعة أشياء متـماثلة مثـل اختيار كرة من صنـدوق به مجمـوعة كرات أو سحب بطاقة من حزمة بطاقات. أو سحب بعض قطع من إنتاج معين لفحصها لمعرفة ما إن كانت سليمة أو معيبة. ونأخذ على سبيل المثال حساب عدد الطرق التي يتم بها سحب كرة واحدة من صندوق به مجموعة كرات متماثلة أو عدد طرق سحب كرتين أو أكثر. وبصورة عامة ما هو عدد الطرق التي يتم بها سحب r كرة بالتتالي من صندوق به n من الكرات المتماثلة؟ ولدراسة هذا الموضوع ينبغي علينا أن نعرف أن السحب يتم بطريقتين مختلفتين الأولى تسمى السحب بإحلال (بإرجاع) والثانية السحب بدون إحلال (بدون إرجاع) وسوف نتناول كل واحدة بالتفصيل كما يلي.

(٧ ـ ٣ ـ ١) السحب بإحلال (بإرجاع)

عند سحب r كرة من صندوق به n من الكرات فإنه إذا سحبت الكرة الأولى تعـاد إلى الصندوق قبل إجراء السحب الثاني وهكذا لباقي السحب. هذه الطريقة تسمى بالسحب بإحلال ولحساب عدد الطرق بهذه الطريقة يتم التالي:

عدد الطرق التي يتم بها السحب الأول يساوي n طريقة، وعدد الطرق التي يتم بها السحب الثـاني يسـاوي n طريقـة أيضًا، وهكذا. وأن عدد الطرق التي يتم بها السحب r يساوي n طريقة أيضًا، وباستخدام القاعدة الأساسية لطرق العد (قاعدة

الضرب) يكون عدد الطرق التي يتم السحب بها r مرة هو:

$$\underbrace{n \times n \times \ldots\ldots \times n}_{\text{r مرة}} = n^r \qquad \ldots\ldots\ldots\ldots \quad (٧ - ٥)$$

مثال (٧ - ١٠)

وعاء به 15 كرة ما هو عدد طرق سحب كرتين بإرجاع .

الحـــل

باستخدام القاعدة السابقة للسحب بإحلال .

حيث $n = 15$, $r = 2$ يكون

$$n^r = 15^2 = 225$$ مرة

(٧ - ٣ - ٢) السحب بدون إحلال (أو بدون إرجاع)

يقال إن السحب يتم بدون إحلال إذا لم يتم إعادة أي سحبة من عناصر العينة المسحوبة إلى المجتمع المسحوب منه العينة. فإذا كان المطلوب سحب عدد r من الكرات مثلاً من صندوق به n من الكرات المتماثلة فإن حساب عدد الطرق يتم في هذه الحالة كما يلي:

عدد طرق سحب الكرة الأولى = n طريقة.

وعدد طرق سحب الكرة الثانية بدون إرجاع الكرة الأولى = $n - 1$ طريقة.

وهكـذا لباقي السحب فيكون عدد طرق سحب الكرة الرائية بدون إرجاع هي $(n - r + 1)$.

وحسب القاعدة الأساسية لطرق العد يكون:

عدد الطرق الكلية لسحب r كرة من n كرة بدون إرجاع يساوي

$$n(n-1)(n-2)\ldots\ldots(n-r+1)$$

أي أن عدد الطرق = ^nP_r

مثال (٧ ـ ١١)

أوجد عدد طرق سحب كرتين بدون إحلال من مثال (٧ ـ ١٠).

الحـــل

عدد الطرق $= {}^{n}P_r$

حيث $n = 15 \quad , \quad r = 2$

طريقة $210 = 15 \times 14 = {}^{15}P_2$

ملحوظة مهمة

تسمى طرق إحلال أو بدون إحلال بالعيّنات المرتبة.

(٧ ـ ٤) التوافيـق (Combinations)

التوفيقة هي كل مجموعة يمكن اختيارها من عدة أشياء مختلفة بأخذها كلها أو بعضها دون مراعاة الترتيب في تلك المجموعة فإذا كان لدينا n من الأشياء المختلفة فإن عدد التوافيق التي يمكن تكوينها بحيث تحتوي كل توفيقة على r من هذه الأشياء سوف يرمز له بالرمز $\binom{n}{r}$ ويمكن إيجاد علاقة بين ${}^{n}P_r$ ، $\binom{n}{r}$ كما يلي:

حيث إن كل توفيقة تحتوي r من الأشياء فإنه يمكن تبديل هذه الأشياء بعدد $r!$ طريقة وبذلك يكون عدد التباديل في جميع التوافيق هو $\binom{n}{r}(r!)$

أي أن
$$ {}^{n}P_r = r!\binom{n}{r} $$

$$ \binom{n}{r} = \frac{{}^{n}P_r}{r!} = \frac{n!}{r!\,(n-r)!} \qquad \text{.........} \qquad (٧ - ٦) $$

مثال (٧ ـ ١٢)

أوجد قيم التوافيق التالية:

$$ \binom{5}{3}, \binom{6}{4}, \binom{4}{0}, \binom{7}{7}, $$

الحـــل

باستخدام العلاقة (٧ ـ ٦) نحصل على :

$$\binom{5}{3} = \frac{5\,!}{3!\ 2!} = 10$$

$$\binom{6}{4} = \frac{6\,!}{4!\ 2!} = 15$$

$$\binom{4}{0} = \frac{4\,!}{0!\ 4!} = 1$$

$$\binom{7}{7} = \frac{7\,!}{7!\ 0!} = 1$$

ويمكن استنتاج بعض العلاقات التالية للتوافيق

$$\binom{n}{r} = \binom{n}{n-r} \qquad (0 \leqslant r \leqslant n) \quad \cdots\cdots\cdots \quad (٧ ـ ٧)$$

مثال (٧ ـ ١٣)

أوجد عدد طرق اختيار حرفين من الحروف A,B,C بدون ترتيب .

الحـــل

طرق اختيار حرفين بدون ترتيب هو:

AB, BC, AC

ويكون عدد الطرق = 3 طرق.

ويمكن الحصول على النتيجة نفسها باستخدام التوافيق $\binom{n}{r}$ حيث $r = 2\ ,\ n = 3$.

$$\binom{3}{2} = \frac{3!}{2!\ 1!} = 3 \ \text{طرق}$$

مثال (٧ ـ ١٤)

أوجد عدد طرق اختيار كرتين من صندوق يحتوي على 15 كرة بدون ترتيب .

الحـــل

طريقة $\binom{15}{2} = \frac{15!}{13!\ 2!} = \frac{15 \times 14}{2} = 105$

(٧ ـ ٤ ـ ١) التباديل داخل أشياء متساوية

نفرض أن لدينا n من الأشياء تحتوي على r من المجموعات كل مجموعة يوجد فيها الأشياء نفسها فإذا كان عدد هذه الأشياء في المجموعات هو $n_1, n_2, \dots, n_r$ (حيث $n_1 + n_2 - \dots + n_r = n$) فسوف نبحث الآن في إيجاد عدد التبديلات الممكنة داخل الأشياء كلها.

عدد طرق اختيار n_1 من الأشياء n هو $\binom{n}{n_1}$.

فيكون الباقي هو $n - n_1$ وبذلك يكون عدد طرق اختيار n_2 من $n - n_1$ هو $\binom{n - n_1}{n_2}$

وهكذا فيكون عدد الطرق المختلفة هو

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\cdots\binom{n-n_1-\dots-n_{r-2}}{n_{r-1}}$$

وباستخدام مفكوك $\binom{n}{r}$ في العلاقة السابقة نحصل على أن عدد التبديلات هو

$$\frac{n!}{n_1!\ n_2!\ \dots\ n_r!}$$

مثال (٧ ـ ١٥)

بكم طريقة يمكن ترتيب حروف كلمة Probability ؟

الحـــل

نجد الحرف (b) مكررًا مرتين والحرف (i) مكررًا مرتين فيكون عدد التراتيب أو التباديل لكلمة $\text{Probability} = \frac{11!}{2!\ 2!}$

$= 9979200$ تبديلة

مثال (٧ ـ ١٦)

موقف مخصص لثماني سيارات مكونة من صف واحد بكم طريقة يمكن وضع 4 سيارات مازدا، 3 سيارات تويوتا، وسيارة واحدة مرسيدس.

الحـــل

طريقة $280 = \frac{8!}{4!\ 3!\ 1!} =$ عدد الطرق الممكنة.

(٧ ـ ٥) تمــارين

(١) أحسب كلاً من:

$$7!\ ,\ 5!\ ,\ \frac{10!}{9!},\ {}^{10}\mathbf{P}_{10}\ ,\ {}^{10}\mathbf{P}_0\ ,\ \binom{7}{2}\ ,\ \binom{7}{4}\ ,\ \binom{8}{0}\ ,\ \binom{8}{3}$$

(٢) إذا كان لدينا الحروف A, B, C, D فبكم طريقة يمكن تكوين ثلاثة حروف في كل من الحالتين التاليتين:

i) بأخذ الترتيب في الاعتبار.

ii) بدون أخذ الترتيب في الاعتبار.

(٣) بكم طريقة يمكت ترتيب حروف كلمة Statistics .

(٤) لدينا الأرقام التالية: 1,2,3,4,5 بفرض عدم السماح بالتكرار أوجد التالي:

i) كم عددًا يمكن تكوينه من ثلاثة أرقام.

ii) كم عددًا يمكن تكوينه بحيث يكون أقل من 300 .

iii) كم عددًا يمكن تكوينه بحيث يكون عددًا زوجيًا.

(٥) بكم طريقـة يمكن أن يجلس أربعـة من الأمريكيين، وثلاثة من الفرنسيين في الحالتين التاليتين:

i) في صف واحد.

ii) حول مائدة مستديرة.

(٦) صندوق به 9 كرات بكم طريقة يمكن اختيار عيّنة مكونة من ثلاث كرات في الحالتين التاليتين :

i) بإحـــلال .

ii) بدون إحـــلال .

(٧) في أحد الإمتحانات لمقرر ١٠١ إحص كان عدد الأسئلة ستة والمطلوب الإجابة على خمسة .

i) بكم طريقة يمكن للطالب أن يختار خمسة أسئلة .

ii) إن كان السؤال الأول والثـاني إجباري بكم طريقة يمكن للطالب إختيار الأسئلة الخمسة؟

(٨) المجموعة A تحتوي على n عنصرًا، باستخدام التوافيق أوجد عدد المجموعات الجزئية المحتوية على r عنصرًا من عناصر هذه المجموعة .

(٩) بكم طريقة يمكن سحب ورقتين من أوراق اللعب في الحالات التالية :

i) بحيث يكون لونهما أحمر .

ii) بحيث يكون لونهما أسود .

iii) واحدة لونها أسود والثانية لونها أحمر .

iv) بحيث يكونان من اللون نفسه .

(١٠) بكم طريقة يمكن إختيار طالبين من بين 5 طلاّب؟

(١١) بكم طريقة يمكن تقسيم 10 طلاّب إلى مجموعتين بحيث تشمل كل مجموعة 5 طلاّب .

(١٢) يوجد بين المدينتين B و A أربعة طرق وبين المدينتين C و B ثلاثة طرق بكم طريقة يمكن لشخص إذا قام من A أن يصل إلى C مارًا بالمدينة B ؟

(١٣) بكم طريقة يمكن ترتيب حروف كلمة سلسبيل؟

(١٤) إذا علم أن أرقام الهواتف الداخلية في جامعة الملك سعود مؤلفة من خمسة أرقام تبدأ دائمًا من اليسار بالرقم 7 فما هو عدد الهواتف التي يمكن تركيبها في الحالتين :

i) التكرار ممكن .

ii) التكرار غير ممكن .

(١٥) تريد مصلحة المرور في إحدى المدن تصميم لوحات معدنية لأرقام السيارات بحيث تحتوي اللوحة على ثلاثة حروف عربية متبوعة بثلاث أرقام عربية بحيث لا يكون الصفر هو الرقم الأخير.

أوجد عدد اللوحات المعدنية المختلفة التي يمكن أن تصدرها هذه المصلحة علمًا بأن عدد الحروف العربية هو 28 حرفًا وعدد الأرقام هو عشرة أرقام .

(١٦) لدينا تسعة طلّاب منهم ستة سعوديون وثلاثة غير سعوديين، نريد اختيار وفد من أربعة طلّاب .

i) ما عدد الطرق الممكنة لاختيار الوفد .

ii) ما عدد الطرق الممكنة لاختيار الوفد إن كان اثنان منهم غير سعوديين .

iii) ما عدد الطرق الممكنة للاختيار إن كان الوفد يحتوي على اثنين على الأكثر من غير السعوديين .

(١٧) اثبت أن : $\sum_{x=0}^{n} \binom{n}{x} = 2^n$

(١٨) أكمل ما يلي :

ا ـ عدد الطرق المختلفة لاختيار لجنة من ثلاثة أشخاص من بين ٩ أشخاص هو ....

ب ـ عدد الطرق المختلفة لترتيب حروف كلمة committe = ...

جـ ـ $^{15}P_2 =$ .... و .... $= \binom{15}{2}$.

د ـ عدد التباديل المختلفة لحروف كلمة ثرثرة هي ....

٨

مبادىء الاحتمالات

Principles of Probability

● مقدمة ● التجربة العشوائية ● فضاء أو فراغ العينة ● الحادثة أو الحدث ● الاحتمال ● حقل سجما (فصل الحوادث) ● مسلمات (بديهيّات) الاحتمالات ● الاحتمال الشرطي ● احتمال الحوادث المستقلة ● نظرية بيز ● تمارين

(٨ - ١) مقدمة

إن كلمة «احتمال» تستخدم كثيراً في حياتنا اليومية. فمثلاً يقال احتمال أن يفوز نادٍ رياضي معيّن بكأس الملك هذا العام، واحتمال أن يسقط المطر اليوم، واحتمال أن يفوز حصان معين بالسباق اليوم، واحتمال أن يكون هذا المصباح غير سليم (محروقًا أو معيبًا)، واحتمال أن تكون قطعة إنتاج معيّنة معيبة... الخ. وكلمة احتمال تكون للتعبير عن حدث بذاته غير مؤكد الحدوث. ويستخدم عدد كبير من الأفراد في الظروف نفسها كلمة احتمال بتعبيرات مختلفة مثل احتمال سقوط المطر اليوم أو احتمال قوي أن يسقط المطر اليوم أو احتمال قوي جدًا بأن يسقط المطر اليوم، حيث تكون درجة الثقة في سقوط المطر مختلفة من شخص لآخر. ومن هنا نشأت الحاجة إلى وضع مقاييس رقمية بدلاً من التعبيرات التي يفهم منها درجة الثقة في وقوع الحدث المُعَبّر عنه. والعلم الذي يبحث في هذه المقاييس وعلاقة بعضها ببعض يسمى علم الاحتمالات. والاحتمالات أساس علم الإحصاء.

وعلم الاحتمالات مثل العلوم الأخرى يبدأ ببعض التعاريف والمسلّمات وسوف نستعرض بعض التعاريف قبل الدخول في مسلّمات الاحتمالات كما يلي.

(٨ - ٢) التجربة العشوائية (Random experiment)

هي التجربة التي تكون جميع نتائجها معلومة مسبقًا ولكن لا يمكن لأحد التنبؤ بحدوث أي من هذه النتائج أو لا (بصفة مؤكدة) فمثلاً عند قذف قطة نقود فإن جميع نتائجها الممكنة هي ظهور الصورة أو ظهور الكتابة ولكن لا يمكن لأحد أن يؤكد قبل رمي العملة أن ما يظهر هو الصورة مثلاً. فعملية الرمي لقطعة النقود تسمى تجربة عشوائية. وكذلك عند فحص أحد المصابيح فإنه يكون إمّا سليمًا (يضيء) أو معيبًا (لا يضيء) فإن عملية الفحص للمصباح تسمى تجربة عشوائية ونتائجها الممكنة هي كون المصباح سالمًا أو كونه معيبًا. وعند رمي حجر نرد مرة واحدة فعملية رمي الحجر تسمى تجربة عشوائية لأن جميع نتائجها هي ظهور الأوجه الستة التي تحمل الأرقام 1, 2, 3, 4, 5, 6 . ويوجد كثير من التجارب العشوائية في الحياة العملية.

(٨ - ٣) فضاء أو فراغ العينة (Sample space)

يعرف فضاء العينة بأنه المجموعة المكونة من جميع النتائج الممكنة للتجربة العشوائية ونرمز لها بالرمز S . كما أن كل نتيجة ممكنة تسمى نقطة عيّنة. إن تجربة قذف قطعة النقود العشوائية تنتهي إلى نتيجتين (عشوائيتين): ظهور صورة (H) وظهور كتابة (T) وبذلك يكون فضاء العينة S هو:

$$S = \{H , T\}$$

وكذلك عند فحص أحد المصابيح فإن نتائج التجربة إما سليم (يضيء) وسوف نرمز له بالرمز (1) أو معيب (محروق) وسوف نرمز له بالرمز (0) فيكون فضاء العينة S هو:

$$S = \{0 , 1\}$$

وكذلك عند رمي حجر النرد مرة واحدة فإن جميع نتائج التجربة العشوائية يكون ظهور الأوجه التي تحمل الأرقام 1, 2, 3, 4, 5, 6 ويكون فضاء العينة S هو:

$$S = \{1, 2, 3, 4, 5, 6\}$$

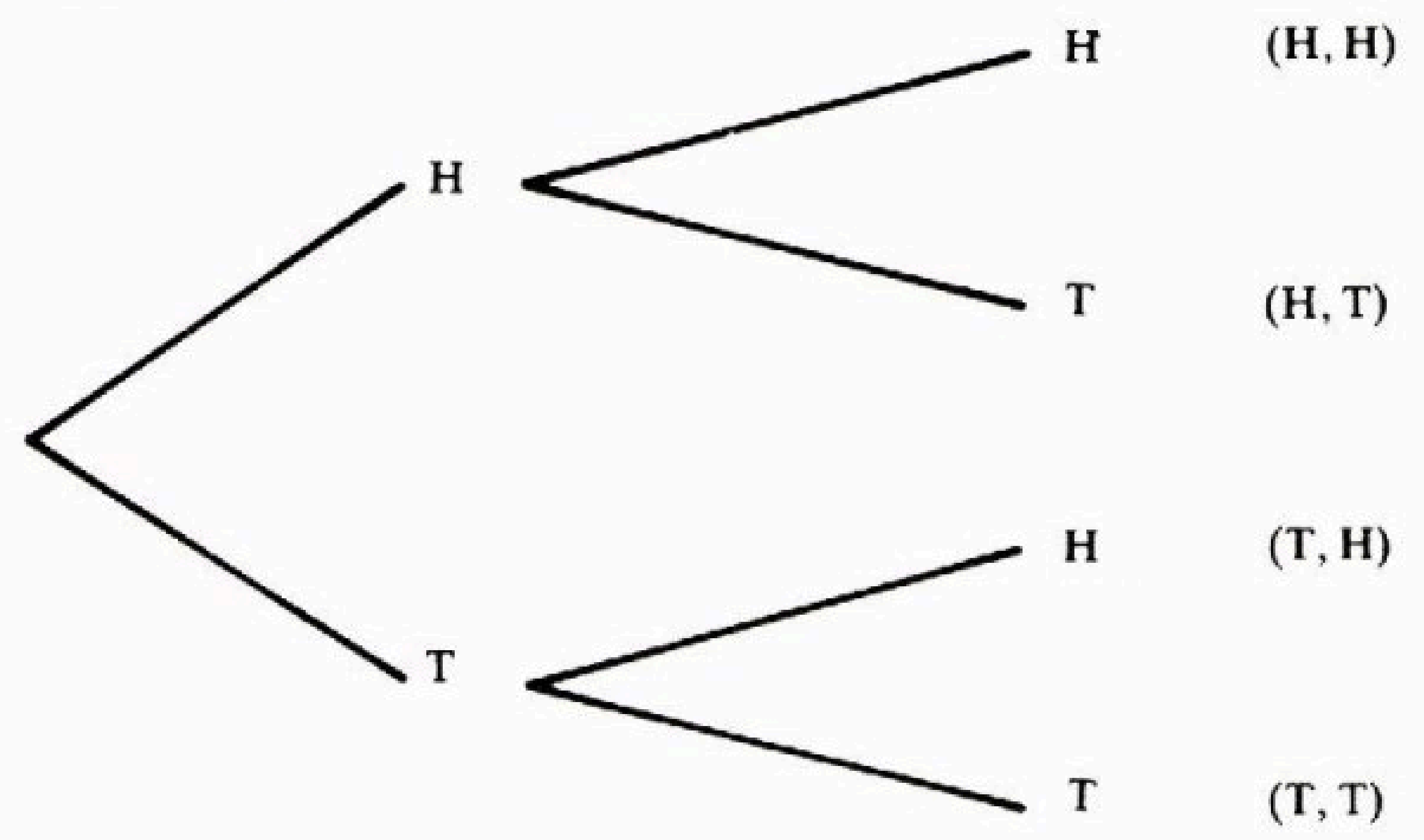
مثال (٨ ـ ١)

أوجد فضاء العيّنة عند قذف قطعة نقود مرتين.

الحـــل

يمكن الحصول على فضاء العيّنة باستخدام الشجرة البيانية، شكل (٨ ـ ١) التالي:

نتيجة الرمية الثانية   نتيجة الرمية الأولى



شكل (٨ ـ ١) يمثل الشجرة البيانية لفراغ العيّنة الناتج من قذف قطعة نقود مرتين

فيكون جميع نتائج التجربة العشوائية وهي قذف قطعة النقود مرتين هو جميع الحالات الناتجة من الفروع المختلفة للشجرة البيانية أي أن فضاء العيّنة S كالتالي:

$$S = \{HH, HT, TH, TT\}$$

وعدد عناصر فضاء العينة S الذي سوف نرمز له بالرمز $n(S)$ يكون مساويًا 4 عناصر ويمكن أيضًا الحصول على فضاء العينة عند قذف قطعة النقود مرتين من الجداء الديكارتي $\{H, T\} \times \{H, T\}$ أي:

$$S = \{HH, HT, TH, TT\}$$

مثال (٨ ـ ٢)

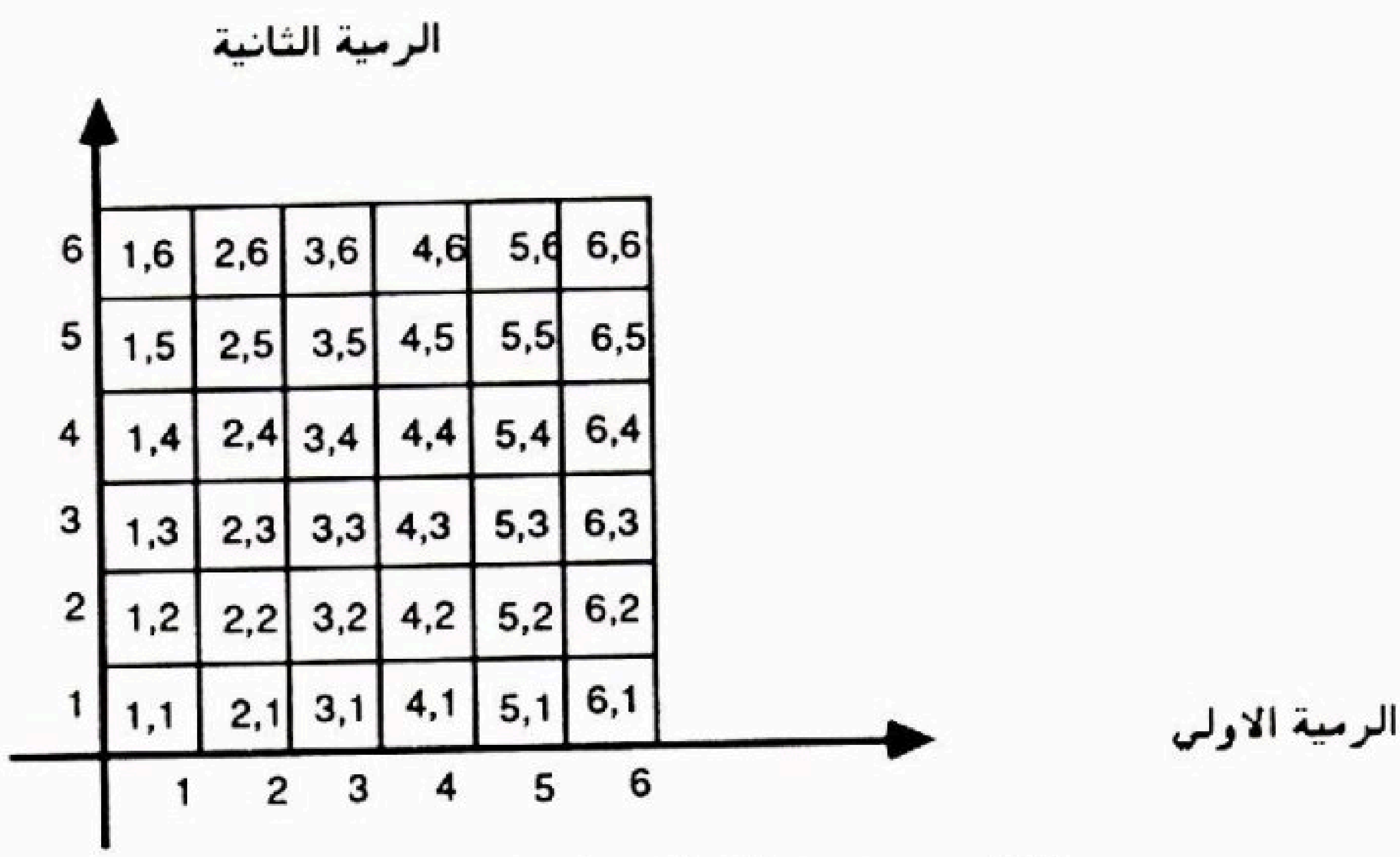
أوجد فضاء العيّنة S عند رمي حجر النرد مرتين متتاليتين وكذلك عدد عناصر فضاء العيّنة: $n(s)$.

الحـــل

يمكن الحصول على فضاء العيّنة S باستخدام الجداء الديكارتي :

$$\{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$$

أو باستخدام شبكة التربيع كما هو موضح في الشكل (٨ ـ ٢) التالي :



شكل (٨ ـ ٢) شبكة التربيع لنتائج رمي حجر نرد مرتين

ويكون فضاء العينة هو:

$$S = \{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2) \ldots, (2,6), \ldots, (6,1), (6,2), \ldots, (6,6)\}$$

ويكون عدد عناصر فضاء العينة S هو:

$$n(S) = 36$$

(٨ ـ ٣ ـ ١) أنواع فضاء العيّنة

يوجد ثلاثة أنواع من فضاء العينة :

فضاء منتهٍ :

وهو الذي يحتوي على عدد محدود من نقاط فضاء العيّنة .

مثال ذلك فضاء العيّنة في الأمثلة السابقة .

فضاء غير منتهٍ قابل للعدّ:

مثال ذلك الفضاء التالي:

$$S = \{0, 1, 2, 3, \ldots\ldots\}$$

أو

S = {x: عدد صحيح غير سالب x}

فضاء غير منتهٍ غير قابل للعدّ:

مثال ذلك:

تسجيل درجة حرارة الجو خلال يوم بإحدى مدن المملكة يكتب:

S = {x: عدد حقيقي x}

وسوف تقتصر دراستنا في هذا الفصل على فضاء العيّنة من النوع الأول أي الفضاء المنتهِ.

(٨ ـ ٤) الحادثة أو الحدث (Event)

تُعرّف الحادثة A بأنها مجموعة جزئية من فضاء العينة S . ويقال إن الحادثة قد وقعت إذا ظهر (عند إجراء التجربة) واحد أو أكثر من النتائج المحتملة للتجربة التي تكوِّن الحادثة A . فإذا كان لدينا فضاء العينة S الناتج من قذف قطعة نقود مرة واحدة فإن $S = \{H,T\}$ ، وإذا كان اهتمامنا بالحادثة التي تمثل ظهور الصورة فقط فإن $A = \{H\}$ وتكون $A \subset S$.

وإذا كان لدينا فضاء العينة $S = \{1, 2, 3, 4, 5, 6\}$ الناتج من رمي حجر النرد مرة واحدة و كانت الحادثة B تمثل ظهور عدد زوجي فإن $B = \{2, 4, 6\}$ وتكون $B \subset S$.

وعلى ذلك فإن الحادثة يمكن أن تحتوي على عنصر واحد من عناصر فضاء العيّنة S وفي هذه الحالة تسمى **بالحادثة الأولية** (البسيطة) مثل حالة ظهور الصورة عند قذف

قطعة النقود مرة واحدة، أو تكون أكثر من عنصر مثل ظهور العدد الزوجي في حالة رمي حجر النرد مرة واحدة. وفضاء العينة S يعتبر مجموعة جزئية من نفسه S . . أي $S \subset S$. وبـذلـك تكـون S عبارة عن حادثة أيضًا ويطلق عليها **الحادثة الأكيدة** . **وكذلك فإن المجموعة الخالية** ϕ هي مجموعة جزئية أيضًا من S أي $\phi \subset S$ وعليه فإن ϕ تمثل أيضًا حادثة ويطلق عليها **الحادثة المستحيلة** .

مثال (٨ ـ ٣)

أحسب الحـوادث التـالية وعدد عناصر كل حادثة بالنسبة للتجربة العشوائية الممثلة في قذف قطعة نقود مرتين :

A = { الحصول على صورة H في الرمية الأولى }

B = { الحصول على كتابة T في الرمية الأولى }

C = { الحصول على صورة واحدة على الأقل }

الحـــل

فضاء العيّنة S هو:

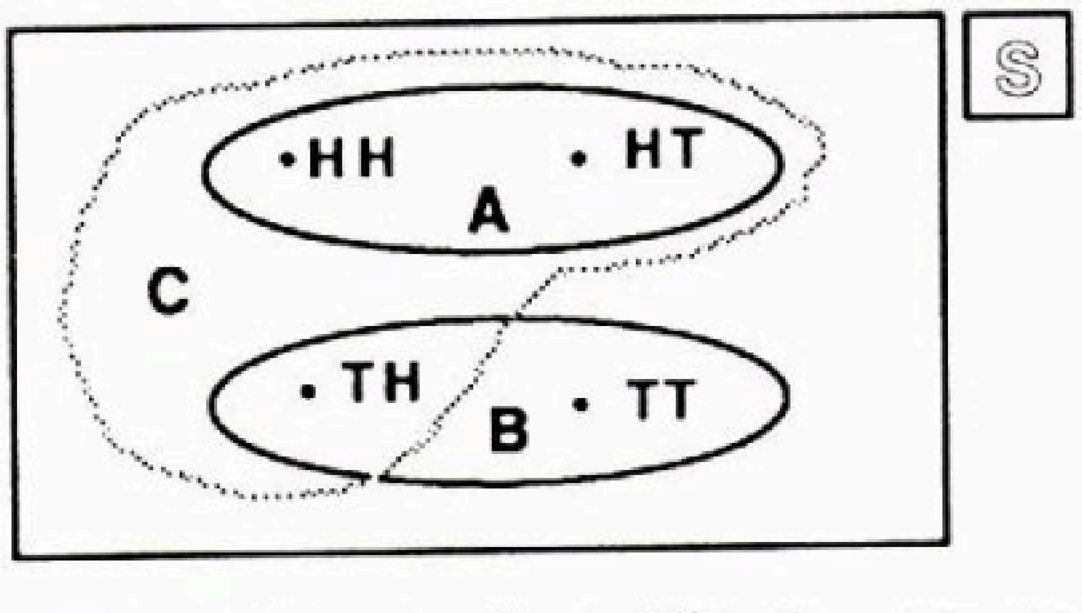
$$S = \{HH, HT, TH, TT\}$$

$$A = \{HH, HT\} \qquad , \; n(A) = 2$$

$$B = \{TH, TT\} \qquad , \; n(B) = 2$$

$$C = \{HH, HT, TH\} \qquad , \; n(C) = 3$$

ونوضح الحل بشكل فن،شكل (٨ ـ ٣) التالي :



شكل (٨ ـ ٣) يمثل شكل فن للحوادث A, B, C وفراغ العيّنة S

مثال (٨ ـ ٤)

أحسب الحوادث التالية وعدد عناصرها بالنسبة لفضاء العيّنة S الناتجة من رمي حجر النرد مرتين متتاليتين، الرمية الأولى (x) والرمية الثانية (y)

$$A = \{(x,y) \ : \ x + y < 4\}$$

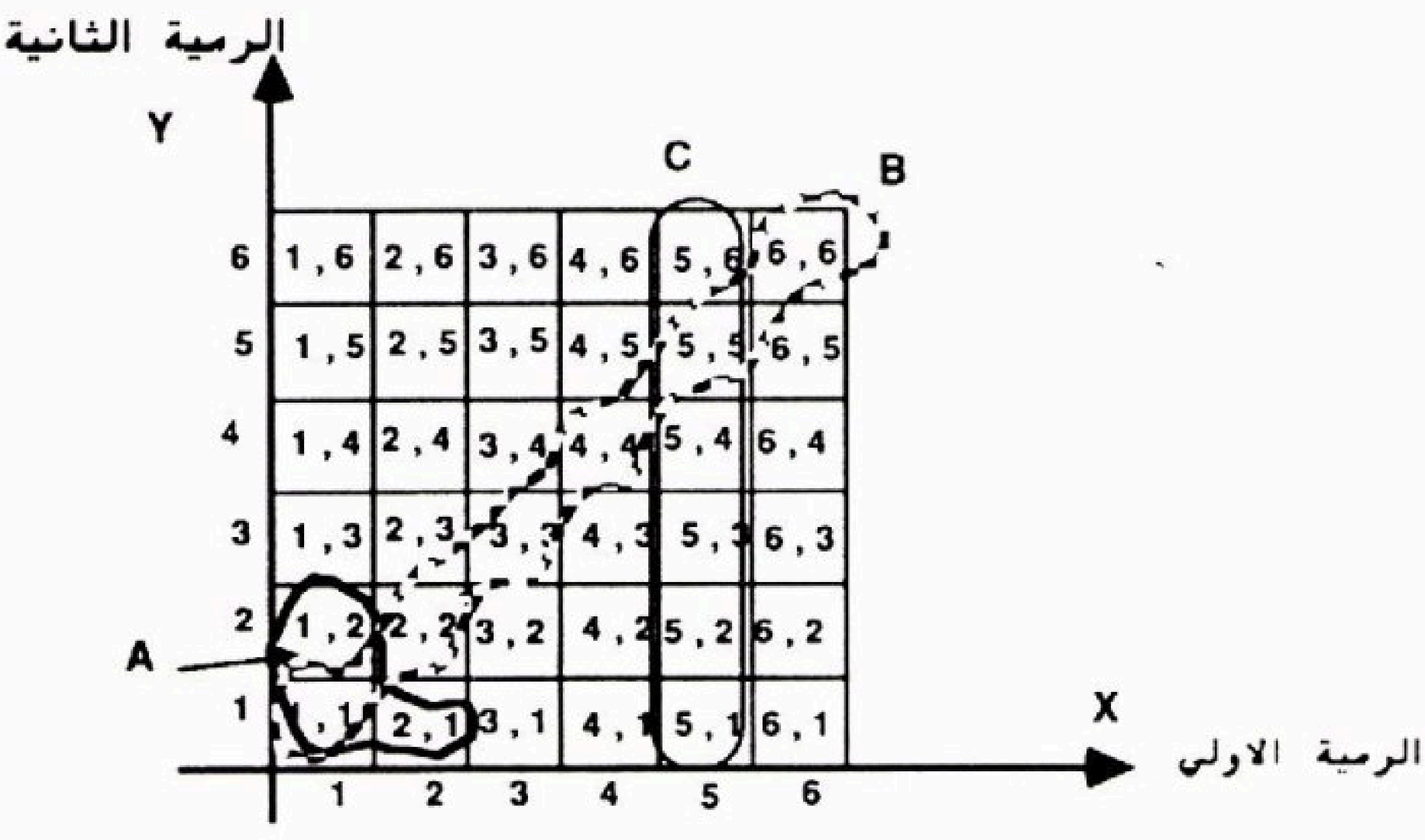
$$B = \{(x,y) \ : \ x = y\}$$

$$C = \{(x,y) \ : \ x = 5\}$$

$$D = \{(x,y) \ : \ x + y = 1\}$$

الحـــل

نكتب فضاء العينة S الناتج من رمي حجر النرد مرتين متتاليتين باستخدام شبكة التربية كما في شكل (٨ ـ ٤) التالي :



شكل (٨ ـ ٤) يمثل شبكة التربيع لرمي حجر نرد مرتين والحوادث A, B, C

ويكون فضاء العيّنة هو:

$$S = \{(1,1) , (1,2) , \ldots , (6,5) , (6,6)\}$$

وتكون الحوادث المطلوبة هي :

$A = \{(1,1) , (1,2) , (2,1)\}$, $n(A) = 3$

$B = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\} \quad , n(B) = 6$

$C = \{(5,1), (5,2), (5,3), (5,4), (5,5), (5,6)\} \quad , n(C) = 6$

$D = \{ \quad \} = \phi \quad , n(D) = 0$

ملاحظة

خلال دراسة الاحتمالات للحوادث نحتاج إلى الحوادث التالية:

(٨ ـ ٤ ـ ١) اتحاد الحادثتين B و A أي $(A \cup B)$

تتكون الحادثة $A \cup B$ من عناصر فضاء العيّنة S والموجودة إما في A أو B أو كلتيهما. ووقوع الحادثة $(A \cup B)$ يعني وقوع إحداهما على الأقل (حدوث إحداهما أو كلتيهما).

(٨ ـ ٤ ـ ٢) تقاطع الحادثتين B و A أي $(A \cap B)$ أو (AB)

تتكون الحادثة $A \cap B$ من عناصر فضاء العينة S الموجودة في كل من الحادثة (A) والحادثة (B) معًا، ووقوع الحادثة $(A \cap B)$ يعني وقوع A ووقوع B معًا. والحادثة (A) والحادثة (B) تسميان حادثتين متنافيتين (منفصلتين) إذا كان $A \cap B = \phi$.

(٨ ـ ٤ ـ ٣) الحادثة المكملة A^c أو $\bar{A}$

الحادثة المكملة A^c تتكون من عناصر فضاء العيّنة S غير الموجودة في الحادثة A .

من التعريف نستنتج أن وقوع الحادثة A^c يقتضي عدم وقوع الحادثة A .

مثال (٨ ـ ٥)

قذفت قطعة نقود متزنة ثلاث مرات:

١ ـ اكتب فضاء العينة S وعدد عناصرها.

٢ ـ اكتب الحوادث التالية وعدد عناصر كل منها:

A : الحادثة الدالة على ظهور صورة في الرمية الأولى.

B : الحادثة الدالة على ظهور صورة واحدة على الأقل.

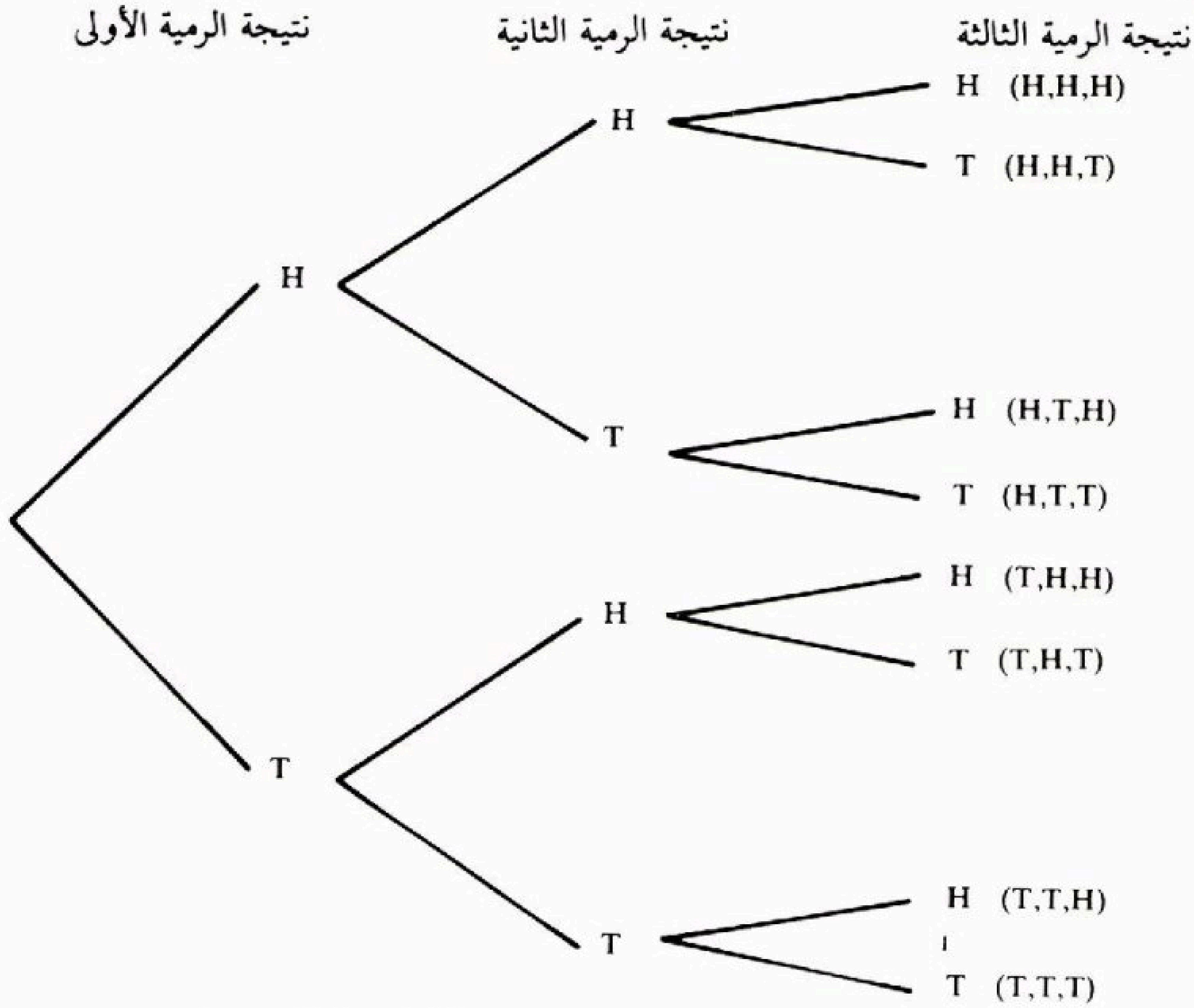
C : الحادثة الدالة على ظهور كتابة في الرمية الأولى وصورة في الرمية الثالثة.

iii) أوجد الحوادث التالية وعناصر كل منها

$$A \cap B \, , \, A \cup C \, , \, A^c \cup B^c \, , \, (A \cap B)^c \, , \, A \cap B^c$$

الحـــل

يمكن الحصول على فضاء العيّنة باستخدام الشجرة البيانية، شكل (٨ ـ ٥) التالي:



شكل (٨ ـ ٥) يمثل الشجرة البيانية لنتائج رمي قطعة عملة ثلاث مرات

i) $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$; $n(S) = 8$

ii) $A = \{HHH, HHT, HTH, HTT\}$; $n(A) = 4$

$B = \{HHH, HHT, HTH, HTT, THH, THT, TTH\}$; $n(B) = 7$

$C = \{THH, TTH\}$; $n(C) = 2$

iii) $A \cap B = \{HHH, HHT, HTH, HTT\}$; $n(A \cap B) = 4$

$$A \cup C = \{HHH, HHT, HTH, HTT, THH, TTH\} \;;\; n(A \cup C) = 6$$

$$A^c = \{THH, THT, TTH, TTT\} \;;\; n(A^c) = 4$$

$$B^c = \{TTT\} \;;\; n(B^c) = 1$$

$$A^c \cup B^c = \{THH, THT, TTH, TTT\} \;;\; n(A^c \cup B^c) = 4$$

$$(A \cap B)^c = S - A \cap B = \{THH, THT, TTH, TTT\} \;;\; n(A \cap B)^c = 4$$

$$A \cap B^c = \phi \;;\; n(A \cap B^c) = 0$$

وقبل الدخول في تعريف الاحتمال نستعرض بعض الحالات التالية.

(٨ ـ ٤ ـ ٤) الحالات المواتية (favourable cases)

هي الحالات التي تؤدي إلى تحقق حادثة معيّنة.

ففي حالة قذف قطعة النقود فإن ظهور الصورة يعتبر حالة مواتية أو مفضلة إذا كان اهتمامنا بالحادثة وهي ظهور الصورة. وفي حالة رمي حجر النرد فإن ظهور الوجه الذي يحمل الرقم 2 أو 4 أو 6 تعتبر حالات مواتية أو مفضّلة إذا كان اهتمامنا بحادثة ظهور رقم زوجي عند إلقاء حجر النرد.

(٨ ـ ٤ ـ ٥) الحالات المتماثلة أو المتساوية الفرصة (equality likely cases)

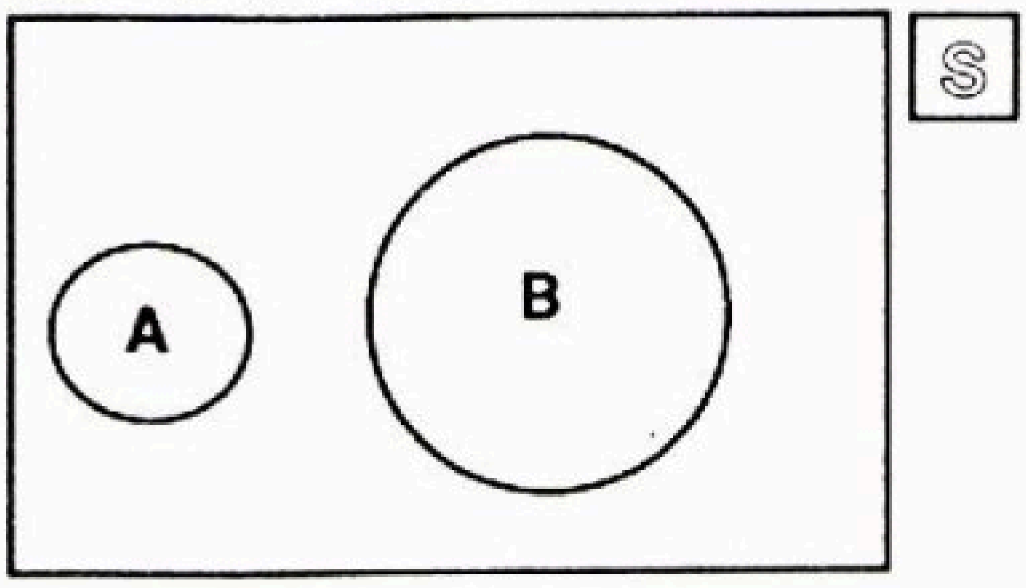
إذا كان لدينـا تجربة عشوائية وكانت جميع نتائج التجربة متساوية الفرصة في الظهور؛ وتكون المصادفة وحدها هي التي تحدد ذلك فإنه يقال لنتائج التجربة متكافئة الفرص ومتماثلة.

مثـال ذلك عند قذف قطعة نقود متزنة فإن فرصة ظهور الصورة (H) مساوية لفرصة ظهور الكتابة (T)، ويعتبر ظهور الصورة والكتابة حالتين متماثلتين. وعند رمي حجر نرد متجانس ومنتظم، فإن ظهور أي وجه من الأوجه الستة يعتبر حالة متماثلة. وعند سحب كرة من صندوق به مجموعة من الكرات متساوية الحجم والكثافة والملمس هي أيضًا حالة متماثلة.

(٨ ـ ٤ ـ ٦) الحالات المتنافية (mutually exclusive events)

إذا استحال وقوع الحادثتين B و A معًا فإنه يقال إن الحادثتين B و A متنافيتان أو مانعتان بعضهما بعضًا ويكون $A \cap B = \phi$ كما موضح بشكل (٨ ـ ٦):
فمثلاً عند قذف قطعة نقود فإن وقوع الحادثة A وهي تمثل ظهور الصورة (H) أي $A = \{H\}$ ينفي وقـوع الحـادثة B والتي تمثل ظهور الكتابة (T) أي أن $B = \{T\}$.
وكذلك عند رمي حجر النرد فإن ظهور الوجه الذي يحمل رقم 1 ينفي ظهور أي وجه آخر من ستة الأوجه الأخرى.

شبكة التربيع النتائج رمي حجر نرد مرتين



شكل (٨ ـ ٦) يمثل شكل فن للحادثتين المتنافيتين A, B

(٨ ـ ٤ ـ ٧) الحوادث الشاملة (exhaustive events)

إذا كان لدينا مجموعة من الحوادث $A_1, A_2, \ldots, A_n$ عند إجراء تجربة عشوائية ما. فإنه يقال لهذه الحوادث السابقة حوادث شاملة إذا كان لا بد من حدوث إحداهما عند إجراء هذه التجربة.

مثال ذلك عند إلقاء حجر النرد فإن الحوادث البسيطة $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$ تعتبر حوادث شاملة.

(٨ ـ ٥) الاحـتـمـال (Probability)

هناك عدة تعاريف للاحتمال نذكر منها.

(٨ ـ ٥ ـ ١) التعريف الكلاسيكي للاحتمال (Classical definition)

إذا كان لدينا تجربة عشوائية جميع نتائجها متماثلة أي أنها متساوية الفرصة في

الظهور وكان فضاء العيّنة لها S يحتوي على عدد منتهٍ من العناصر n(S) وكان لدينا حادثة A تحتوي على n(A) من العناصر المتماثلة فإن الاحتمال الكلاسيكي للحادثة A ويرمز له P(A) (ويقرأ احتمال A) ويحسب كالتالي :

$$P(A) = \frac{n(A)}{n(S)} \qquad \text{(٨ - ١)}$$

مثال (٨ - ٦)

إذا قذفت قطعة نقود متزنة مرتين أوجد احتمال الحصول على الصورة مرتين .

الحـــل

فضاء العينة S هو:

$$S = \{HH, HT, TH, TT\} \quad ; \quad n(S) = 4$$

الحادثة A تمثل ظهور الصورة مرتين هي :

$$A = \{HH\} \quad , \quad n(A) = 1$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{1}{4}$$

(٨ - ٥ - ٢) الاحتمال النسبي

عيب التعريف الكلاسيكي أنه مبني على تساوي الفرص في الظهور لكل نتيجة من جميع النتائج الممكنة للتجربة العشوائية ، وهناك تعريف آخر لاحتمال حدوث حادثة ما هو ما يسمى بالاحتمال التجريبي (النسبي) وهو مبني على فكرة التكرار النسبي فإذا ما كررنا تجربة عشوائية n من المرات وكان عدد مرات ظهور الحادثة A هو r(A) فإن الاحتمال P(A) يعطى بالعلاقة :

$$P(A) = \lim_{n \to \infty} \frac{r(A)}{n} \qquad \text{(٨ - ٢)}$$

فمثلاً عند قذف قطعة عملة متزنة عدد n من المرات وكانت عدد مرات حدوث الحادثة A (الحادثة A تمثل الصورة H) هو r(A) فإن الاحتمال P(A) هو

$$P(A) = \lim_{n \to \infty} \frac{r(A)}{n} = \frac{1}{2}$$

وعلى الرغم من أن هذا التعريف مفيد إلا أننا لا نستطيع أن نؤكد أننا سوف نحصل على النسبة نفسها إذا أجريت التجربة n من المرات في وقت آخر. وكذلك له صعوباته من وجهة النظر الرياضية، فقد تكون هذه النهاية غير موجودة بالفعل. لذلك فقد ظهر التعريف الرياضي للاحتمال الذي يعتمد على عدة فرضيات تتناسب مع فكرتنا الإدراكية لمعنى الاحتمال وهذه الفرضيات هي ما تسمى بمسلمات الاحتمال وهي التي سوف نتعرض لها فيما بعد.

(٨ ـ ٦) حقل سجما (فصل الحوادث)

حقل سجما هو عبارة عن مجموعة عناصرها تتكون من بعض المجموعات الجزئية التي يمكن تكوينها من فراغ العيّنة S . ويرمز لحقل سجما بالرمز ξ (لأننا سوف نقصر دراستنا على فضاء العيّنة المنته فإن هذا التعريف كافٍ لأغراضنا).

مثال (٨ ـ ٧)

ا) أوجد مجموعة القوى P_A المرتبطة برمي حجر نرد مرة واحدة.

ب) أذكر مثالاً لأحد حقول سجما المرتبط برمي حجر نرد مرة واحدة.

الحـــل

ا)

$$S = \{1, 2, 3, 4, 5, 6\}$$

ومن فراغ العيّنة هذا الذي يحتوي على جميع الحوادث الأولية يمكن تكوين المجموعات الجزئية التالية التي تعتبر عناصر مجموعة القوى P_A . ويلاحظ أن هذه المجموعات الجزئية عبارة عن حوادث يطلق عليها اسم الحوادث العشوائية وتتكون من:

$1 = \binom{6}{0}$ وهي الحادثة المستحيلة.

$6 = \binom{6}{1}$ وهي الحوادث الأولية (ذات العنصر الواحد).

$15 = \binom{6}{2}$ وهي الحوادث التي تحتوي على عنصرين.

$20 = \binom{6}{3}$ وهي الحوادث التي تحتوي على ثلاثة عناصر.

$15 = \binom{6}{4}$ وهي الحوادث التي تحتوي على أربعة عناصر.

$6 = \binom{6}{5}$ وهي الحوادث التي تحتوي على خمسة عناصر.

$1 = \binom{6}{6}$ وهي الحوادث التي تحتوي على ستة عناصر.

نلاحظ أن عدد عناصر مجموعة القوى هو 64 مجموعة جزئية (2^6) بوجه عام إذا كان فضاء العيّنة S يحتوي n من الحوادث الأولية فإن مجموعة القوى P_A يحتوي على (2^n) من الحوادث العشوائية.

ب)

$$\xi = \{\phi, S, \{1\}, \{2, 3, 4, 5, 6\}\}$$

ملاحظـة

إذا كان فضاء العيّنة S غير منتهٍ وقابل للعدّ فإن حقل سجما ξ يجب أن يحقق شرط الانغلاق لكل من المتممة والاتحاد لجميع الحوادث العشوائية التي هي عناصر في ξ ، ويحتوي أيضًا على ϕ و S .

(٨ ـ ٧) مسلّمات (بديهيّات) الاحتمالات (Probability axioms)

إذا كان لدينا تجربة عشوائية لها فضاء عيّنة S وكان ξ حقل سجما، وأن P دالة حقيقية معرفة على ξ وكان لدينا الحادثة A فإن P تسمى دالة احتمال ويسمى العدد P(A) باحتمال الحادثة A إذا تحققت المسلّمات التالية:

I ـ لكل حادثة A فإن $P(A) \geqslant 0$

II ـ احتمال فراغ العينة S يكون مساويًا 1 أي $P(S) = 1$

III ـ لأي متابعة من الحوادث $A_1, A_2, \ldots$ المتنافية مثنى مثنى

أي أن :

$$(A_i \cap A_j = \phi \,;\, i,j = 1, 2, \ldots, i \neq j)$$

فإن $P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$

أي أن $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

وهذه المسلّمة تعني إذا كانت الحوادث متنافية فاحتمال اتحادها يساوي مجموع احتمالاتها.

(٨ ـ ٧ ـ ١) بعض النتائج الأساسية لمسلّمات الاحتمال

نتيجة (٨ ـ ١)

$$P(\phi) = 0$$

البرهان

نأخذ المتوالية $A_1, A_2, \ldots$

حيث $A_1 = S$ و $A_i = \phi$ لكل $i > 1$

وعليه فإن : $$S = \bigcup_{i=1}^{\infty} A_i$$

وبما أن هذه الحوادث متنافية مثنى مثنى فإنه حسب المسلمة III يكون :

$$P(S) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

$$P(S) = P(S) + \sum_{i=2}^{\infty} P(\phi)$$

$$\sum_{i=2}^{\infty} P(\phi) = P(S) - P(S) = 0$$

أي أن $$P(\phi) = 0$$

نتيجة (٨ ـ ٢)

إذا كانت $A_1, A_2, \dots, A_n$ حوادث متنافية مثنى مثنى بحيث $A_i \cap A_j = \phi$ لكل $i \neq j$ فإن:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

البرهان:

نأخذ المتوالية $A_1, A_2, \dots, \dots$

بحيث $A_i = \phi$ لكل $i > n$ نلاحظ أن:

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{\infty} A_i$$

وباستخدام المسلّمة III

$$P(\bigcup_{i=1}^{n} A_i) = P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

$$= \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(\phi)$$

$$= \sum_{i=1}^{n} P(A_i)$$

وهو المطلوب

ملاحظة:

عندما $n = 2$ في النتيجة (٨ ـ ٢) يكون:

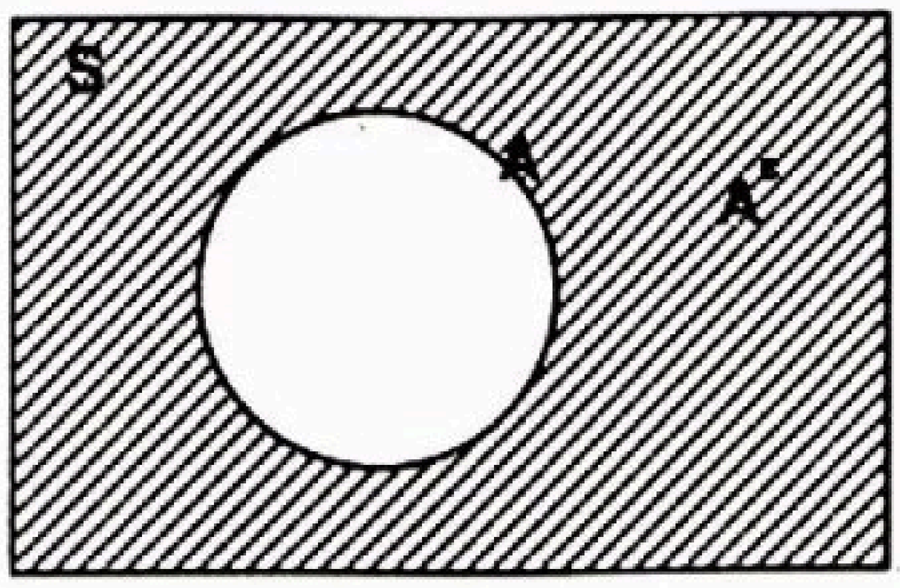
$$P(\sum_{i=1}^{2} A_i) = P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

نتيجة (٨ ـ ٣)

$$P(A^c) = 1 - P(A)$$

البرهان

نوضح الحادثتين A, A^c بشكل فن، شكل (٨ ـ ٧) التالي:



شكل (٨ ـ ٧) يمثل شكل فن الحادثتين A, A^c

الحادثتان A^c, A تحققان التالي :

$$A \cap A^c = \phi \ ; \ A \cup A^c = S$$

ومن المسلّمة III فإن

$$P(A \cup A^c) = P(S)$$

$$P(A) + P(A^c) = P(S)$$

ومن المسلّمة II

$$P(A) + P(A^c) = 1$$

أي أن :

$$P(A^c) = 1 - P(A)$$

وهو المطلوب .

مثال (٨ ـ ٨)

إذا كان احتمال نجاح محمد $\frac{5}{8}$ فما هو احتمال رسوبه .

الحـــل

$A = \{$ نجاح محمد $\}$, $P(A) = \frac{5}{8}$

$A^c = \{$ رسوب محمد $\}$, $P(A^c) = ?$

$$P(A^c) = 1 - P(A) = 1 - \frac{5}{8} = \frac{3}{8}$$

وهو المطلوب .

نتيجة (٨ ـ ٤)

لأي حادثتين B و A نجد أن :

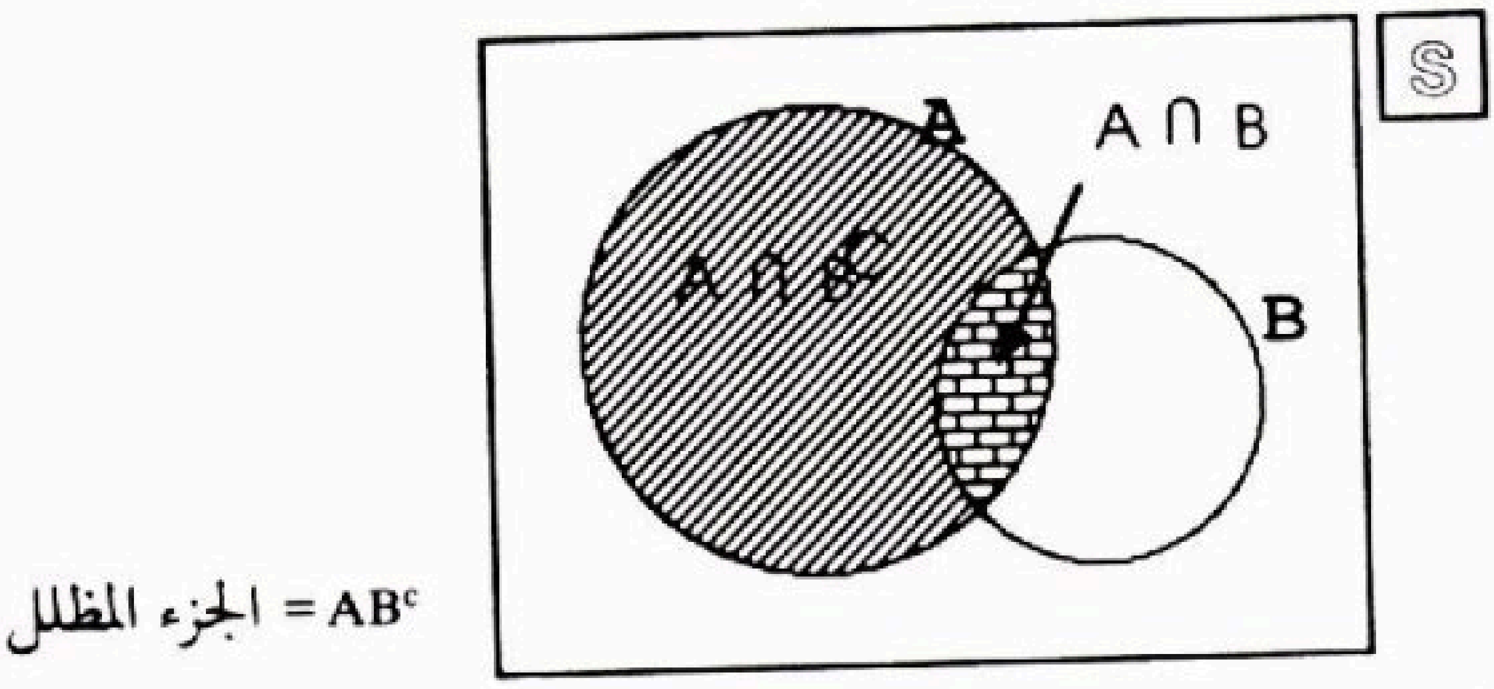
$$P(A \cap B^c) = P(A) - P(A\,B)$$

أو

$$P(A) = P(A\,B) + P(A\,B^c)$$

البرهان

نوضح الحوادث A, B, S بشكل فن شكل (٨ ـ ٨) التالي :



شكل (٨ ـ ٨) يمثل شكل فن للحوادث $A \cap B, A \cap B^c$

الحادثتان AB^c و AB متنافيتان فيكون :

$$(AB)\,(AB^c) = \phi$$

$$A = AB \cup AB^c$$

ومن المسلمة III

$$P(AB) + P(AB^c) = P(A)$$

أي أن

$$P(AB^c) = P(A) - P(AB)$$

وهو المطلوب .

مثال (٨ ـ ٩)

إذا كان احتمال نجاح محمد هو $\frac{5}{8}$ واحتمال نجاح أحمد ومحمد هو $\frac{1}{8}$ فأحسب احتمال نجاح محمد ورسوب أحمد .

الحــل

نفرض أن الحادثتين B و A هما:

$A =$ { نجاح محمد } , $B =$ { نجاح أحمد }

$P(A) = \frac{5}{8}$, $P(AB) = \frac{1}{8}$

$AB^c =$ { نجاح محمد ورسوب أحمد } , $P(AB^c) = ?$

$$\because\ P(AB^c) = P(A) - P(AB)$$

$$\therefore\ P(AB^c) = \frac{5}{8} - \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

وهو المطلوب

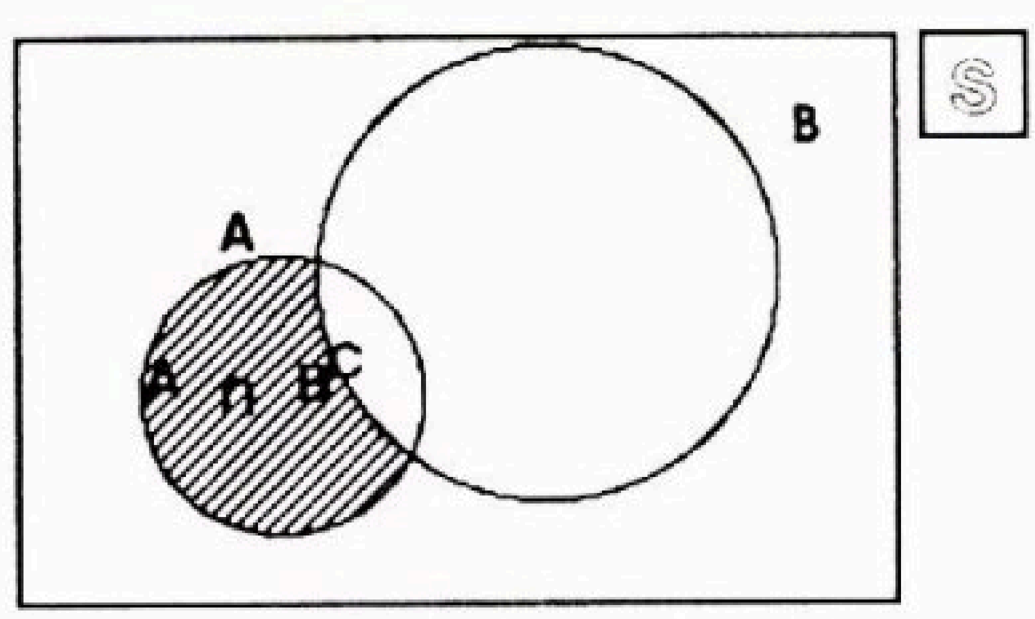
نتيجة (٨ ـ ٥)

لأي حادثتين A, B فإن

$$P(A \cup B) = P(A) + P(B) - P(A\,B)$$

البرهان

شكل فن،شكل (٨ ـ ٩) للحوادث A, B, S التالي:



شكل (٨ ـ ٩) يمثل شكل فن للحوادث A, B, AB^c

من شكل فن الحادثتين B و AB^c متنافيتان أي أن:

$$(AB^c)(B) = \phi$$

وأن

$$A \cup B = (AB^c) \cup B$$

بتطبيق المسلّمة (III)

$$P(A\cup B) = P(A\cap B^c) + P(B)$$

وبالتعويض عن الاحتمال $P(A\,B^c)$ من نتيجة (٨ ـ ٤) نحصل على

$$P(A\cup B) = P(A) - P(AB) + P(B)$$

أي أن

$$P(A\cup B) = P(A) + P(B) - P(AB)$$

وهو المطلوب.

مثال (٨ ـ ١٠)

إذا كان احتمال نجاح محمد $\frac{1}{4}$ واحتمال رسوب أحمد $\frac{1}{3}$ واحتمال نجاح محمد وأحمد $\frac{1}{6}$ أوجد احتمال نجاح أحدهما على الأقل.

الحـــل

نكتب الحوادث التالية:

$P(A) = \frac{1}{4}$, $A = \{$ نجاح محمد $\}$

$P(A^c) = \frac{1}{3}$, $B^c = \{$ رسوب أحمد $\}$, $B = \{$ نجاح أحمد $\}$

$P(AB) = \frac{1}{6}$, $AB = \{$ نجاح محمد وأحمد $\}$

$P(A\cup B) = ?$, $A\cup B = \{$ نجاح محمد أو أحمد أو كليهما $\}$

$$P(B) = 1 - P(B^c) = 1 - \frac{1}{3} = \frac{2}{3}$$

$$P(A\cup B) = P(A) + P(B) - P(AB)$$

$$\therefore\ P(A\cup B) = \frac{1}{4} + \frac{2}{3} - \frac{1}{6} = \frac{3+8-2}{12} = \frac{9}{12}$$

نتيجة (٨ ـ ٦)

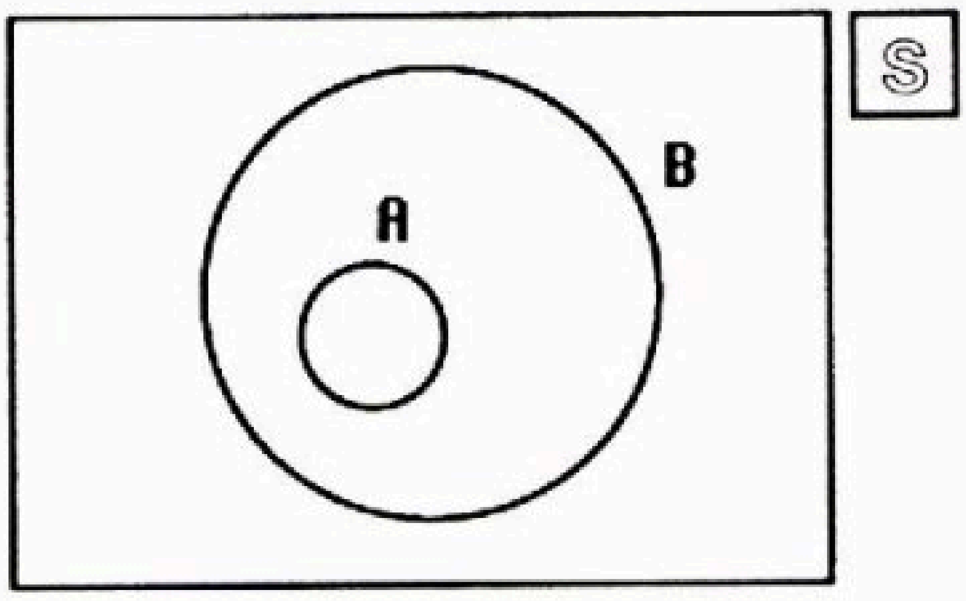
إذا كان $A \subset B$

فإن

$$P(A) \leqslant P(B)$$

البرهان

تمثل الحوادث A, B, S بشكل فن شكل (٨ ـ ١٠) التالي:



شكل (٨ ـ ١٠) يمثل شكل فن للحادثة $A \subset B$

يمكن التعبير عن الحادثة B كالتالي:

$$B = A \cup (A^c B)$$

حيث

$$A \cap (A^c \cap B) = \phi$$

$$P(B) = P(A \cup (A^c B))$$

ومن المسلّمة III

$$P(B) = P(A) + P(A^c B)$$

$$P(B) - P(A) = P(A^c B)$$

$$P(B) - P(A) \geqslant 0$$

حيث إنه لأي حادثة $A^c B$ يكون

$$P(A^c B) \geqslant 0$$

وعليه فإن :

$$P(A) \leqslant P(B)$$

وهو المطلوب .

(٨ - ٧ - ٢) أمثلة متنوعة

مثال (٨ - ١١)

قذفت قطعة نقود متزنة مرتين .

أحسب الاحتمالات التالية :

$$P(A), P(B), P(C), P(AB), P(A\cup B), P(\overline{A\cup B}), P(B\bar{C}), P(\bar{B}\cup\bar{C})$$

حيث A, B, C كالتالي :

A الحادث الدال على ظهور صورة في الرمية الأولى .

B الحادث الدال على ظهور كتابة في الرمية الأولى .

C الحادث الدال على ظهور صورة واحدة على الأقل .

الحـــل

نكتب فراغ العينة S وهو

$$S = \{HH, HT, TH, TT\}$$

$$n(S) = 4$$

ونكتب الحوادث A, B, C وكذلك احتمال كل حادثة على حده كالتالي :

$$A = \{HH, HT\} \quad , \quad n(A) = 2$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{2}{4} = \frac{1}{2}$$

$$B = \{TH, TT\} \ , \ n(B) = 2$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{2}{4} = \frac{1}{2}$$

$$C = \{HH, HT, TH\} \ , \ n(C) = 3$$

$$P(C) = \frac{n(C)}{n(S)} = \frac{3}{4}$$

$$AB = \phi \quad , \quad n(AB) = 0$$

$$P(AB) = \frac{n(AB)}{n(S)} = \frac{0}{4} = 0$$

$$P(A \cup B) = P(A) + P(B) - P(A\,B)$$

$$P(A \cup B) = \frac{1}{2} + \frac{1}{2} - 0 = 1$$

$$P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 1 = 0$$

$$BC = \{TH\} \quad , \quad P(BC) = \frac{n(BC)}{n(S)} = \frac{1}{4}$$

$$P(B\bar{C}) = P(B) - P(BC)$$

$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$P(BC) = \frac{1}{4}$$

$$P(\bar{B} \cup \bar{C}) = P(\overline{B\,C}) = 1 - P(B\,C)$$

$$= 1 - \frac{1}{4} = \frac{3}{4}$$

مثال (٨ ـ ١٢)

إذا كان لدينا حادثتان A, B بحيث

$$P(A) = \frac{1}{5} \; , \; P(B) = \frac{2}{5} \; , \; P(A \cup B) = \frac{1}{2}$$

أحسب

$$P(A\,B) \; , P(A\,B^c) \; , P(A^c\,B^c)$$

الحـــل

$$P(A \cup B) = P(A) + P(B) - P(A\,B)$$

$$\frac{1}{2} = \frac{1}{5} + \frac{2}{5} - P(A\,B)$$

$$P(A\,B) = \frac{3}{5} - \frac{1}{2} = \frac{1}{10}$$

$$P(A\,B^c) = P(A) - P(A\,B)$$

$$= \frac{1}{5} - \frac{1}{10} = \frac{1}{10}$$

$$P(A^c\,B^c) = P(A \cup B)^c = 1 - P(A \cup B)$$

$$= 1 - \frac{1}{2} = \frac{1}{2}$$

مثال (٨ ـ ١٣)

اختير رقم من الأرقام الصحيحة من 1 إلى 50 بطريقة عشوائية (بحيث إن احتمال ظهور الأرقام متساوٍ).

أحسب احتمال أن يكون الرقم هو 4 أو مضاعفاتها.

الحـــل

فراغ العينة S هو

$$S = \{1, 2, 3, \ldots, 50\}, \quad n(S) = 50$$

A = { الرقم 4 أو مضاعفات الـ 4 }

$$A = \{4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48\}, \quad n(A) = 12$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{12}{50} = \frac{6}{25}$$

مثال (٨ ـ ١٤)

إذا اخترنا ورقتين من أوراق اللعب ما هو احتمال أن يكون لونهما أسود؟

الحـــل

عدد أوراق اللعب الكلية = 52 ورقة.

عدد الأوراق التي لونها أسود = 26 ورقة.

الحادثة A هي أن الورقتين المسحوبتين لونهما أسود.

عدد حالات النجاح في الحصول على ورقتين لونهما أسود $\binom{26}{2} = n(A)$

عدد الحالات الكلية $\binom{52}{2} = n(S)$

$$P(A) = \frac{n(A)}{n(S)} = \frac{\binom{22}{2}}{\binom{52}{2}} = \frac{26x25}{52x51} = 0.245$$

(٨ - ٨) الاحتمال الشرطي (Conditional probability)

عندما نتحدث عن احتمال وقوع الحدث A فإننا حقيقة نعني تعريف احتمال وقوع A منسوبًا إلى فضاء العينة محددة (في هذه الحالة S). لكن عند تطبيق نظرية الاحتمالات ترد أوضاع نحتاج فيها إلى إيجاد احتمال الحدث A علمًا بوقوع الحدث B .

فمثلًا:

١) نفرض أن صندوقًا يحتوي على 100 مصباح، 5 منها معيبة ثم سحبنا مصباحين بدون إرجاع. أوجد احتمال أن يكون المصباح المسحوب ثانيًا معيبًا علمًا بأن المصباح المسحوب أولًا كان سليمًا.

٢) أوجد احتمال أن يستمر المصباح صالحًا لمدة 100 ساعة علمًا بأنه ظل صالحًا لمدة 24 ساعة.

يفهم مما سبق أننا نريد إيجاد احتمال الحدث A ، منسوبًا لفضاء عيّنة جديد هو فضاء العيّنة المكون من عناصر الحدث B ، وأن الحدث (B) هو فضاء العيّنة الجديد ويسمى بالفضاء المختزل. $(B \subset S)$ ومرادنا هو احتمال حدوث A بشرط حدوث B .

ونوضح الاحتمال الشرطي باعتبار المثال العددي التالي:

إذا صنفنا طلّاب شعبتين لمقرر ١٠١ إحص حسب تخصص الإحصاء ومن

خارج تخصص الإحصاء وكانت نتيجة التصنيف كما هو موضح بالجدول التالي :

الشعبة / التخصص	الشعبة الأولى	الشعبة الثانية	المجموع
تخصص إحصاء	30	25	55
من خارج تخصص الإحصاء	10	15	25
المجمـــوع	40	40	80

فإذا اخترنا طالبًا بطريقة عشوائية فإننا نلاحظ أن :

i) احتمال أن يكون الطالب من الشعبة الأولى وتخصصه إحصاء يساوي $\frac{30}{80}$ ..

ii) فإذا كان علمنـا أن الطالب المختار من الشعبة الأولى فإن احتمال أن يكون تخصصه إحصاء يساوي $\frac{30}{40}$.

والاحتمال (i) ناتج من أن عدد الحالات المواتية 30 ، وعـدد الحالات الممكنـة هي 80 ، والاحتمال (ii) هو احتمال أن يكون الطالب من تخصص الإحصاء إذا علم أنه من الشعبة الأولى ويسمى احتمالاً شرطيًا. فإذا رمزنا للحادثة A بأن الطالب المختار من تخصص الإحصاء، وللحادثة B بأن الطالب المختار من الشعبة الأولى، فإن احتمال (i) هو P(AB) وأن الاحتمال (ii) هو P(A/B) . ويقرأ احتمال حدوث A إذا حدث B (أو احتمال حدوث A بشرط حدوث B)، ونلاحظ أن :

$$P(A/B) = \frac{30}{40} = \frac{30}{80} / \frac{40}{80} = \frac{P(AB)}{P(B)}$$

من التعريف النسبي للاحتمال نجد أن P(A/B) يمثل نسبة العناصر في فضاء العيّنة S والمشتركة بين الحدثين A,B منسوبة لعناصر الحدث B أي أن :

$$P(A/B) = \frac{n(AB)}{n(B)} = \frac{n(AB) / n(S)}{n(B) \ / n(S)}$$

حيث $n(S)$ عدد العناصر فضاء العينة S .

وينتج من هذا أن :

$$p(A/B) = \frac{P(A\,B)}{P(B)} \qquad \text{.......................} \qquad (٨ - ٣)$$

بحيث إن $P(B) \neq 0$

مثال (٨ - ١٥)

إذا كان احتمال أن ينجح محمد هو $\frac{1}{3}$ واحتمال أن ينجح محمد وأحمد هو $\frac{1}{4}$.

أوجد احتمال نجاح أحمد إذا علم أن محمدًا قد نجح .

الحـــل

نكون الحوادث التالية :

B = { نجاح محمد } ، $P(B) = \frac{1}{3}$

A = { نجاح أحمد } ،

AB = { نجاح محمد وأحمد } ، $P(AB) = \frac{1}{4}$

A/B = { نجاح أحمد بشرط نجاح محمد }

وبذلك يكون المطلوب هو حساب : $P(A/B)$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{3}} = \frac{3}{4}$$

(٨ - ٩) احتمال الحوادث المستقلّة (Independent events)

يقال الحادثتان A,B مستقلتان إذا كان حدوث إحداهما لا يؤثر في حدوث الأخرى أو عدم حدوثه . فمثلاً عند قذف قطعة نقود متزنة مرتين فإن ظهور الصورة في الرمية الأولى (A) لا يؤثر على ظهور الصورة في الرمية الثانية (B) ، ويقال إن ظهور الصورة في الرمية الأولى وفي الرمية الثانية حادثتان مستقلتان ويعبر عن ذلك كما يلي :

$$P(B/A) = P(B) \qquad \text{.....................} \qquad (٨ - ٤)$$

ومن تعريف الاحتمال الشرطي نحصل على

$$\frac{P(A\,B)}{P(A)} = P(B)$$

$$P(A\,B) = P(A)\,.\,P(B) \qquad \text{.....................} \qquad (٨ - ٥)$$

ومن ذلك يمكن القول إن الحادثتين A,B مستقلتان إذا تحقق أحد الشروط التالية:

i) $P(A/B) = P(A)$

ii) $P(B/A) = P(B)$

iii) $P(A\,B) = P(A)\,.\,P(B)$

مثال (٨ - ١٦)

إذا كان احتمال أن يصيب محمد هدفًا معينًا هو $\frac{3}{4}$ واحتمال أن يصيب أحمد الهدف نفسه هو $\frac{1}{3}$. فأوجد احتمالات الحوادث التالية:

١) أن لا يصيب محمد الهدف.

٢) أن لا يصيب أحمد الهدف.

٣) أن يصيبا الهدف معًا.

٤) على الأقل يصيب الهدف أحدهما.

٥) محمد يصيب الهدف وأحمد لا يصيبه.

٦) أن لا يصيبا الهدف معًا.

الحــل

A= { يصيب محمد الهدف } , $P(A) = \frac{3}{4}$

B= { يصيب أحمد الهدف } , $P(A) = \frac{1}{3}$

١) A^c = { لا يصيب محمد الهدف } , $P(A^c) = 1 - P(A) = \frac{1}{4}$

٢) B^c = { لا يصيب أحمد الهدف } , $P(B^c) = 1 - P(B) = \frac{2}{3}$

٣) $AB = \{$ يصيبا الهدف معًا $\}$

وحيث إن إصابة محمد مستقلة عن إصابة أحمد فإن :

$$P(A\cap B) = P(A) \ . \ P(B) = \frac{3}{4} . \frac{1}{3} = \frac{1}{4}$$

(٤) $A\cup B = \{$ على الأقل يصيب الهدف أحدهما $\}$

$$P(A\cup B) = P(A) + P(B) - P(A\,B)$$

$$P(A\cup B)\ \frac{3}{4} + \frac{1}{3} - \frac{1}{4} = \frac{5}{6}$$

٥) $A\,B^c = \{$ محمد يصيب وأحمد لا يصيب $\}$

$$P(A\,B^c) = P(A) - P(A\,B)$$

$$P(A\,B^c) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

٦) $(\overline{A\cap B}) = \{$ أن لا يصيبا معًا الهدف $\}$

$$P(\overline{A\cap B}) = 1 - P(A''B)$$

$$= 1 - \frac{1}{4} = \frac{3}{4}$$

مثال (٨ ـ ١٧)

إذا كانت الحادثتان A, B مستقلتين أثبت أن الحادثتين A, B^c مستقلتان .

أي أثبت أن :

$$P(A\,B^c) = P(A) \ . \ P(B^c)$$

الحـــل

$$P(A\,B^c) = P(A) - P(A\cap B)$$

ولكن A, B مستقلتان فإن

$$P(A\cap B) = P(A) \ . \ P(B)$$

ومن ذلك نحصل على :

$$P(A\cap B^c) = P(A) - P(A) \ . \ P(B)$$

$$= P(A)\ [1-P(B)]$$
$$= P(A)\ .\ P(B^c)$$

وهو المطلوب .

مثال (٨ ـ ١٨)

إذا كانت A, B مستقلتين فأثبت أن الحادثتين A^c, B^c مستقلتان .
أي أثبت أن :

$$P(A^c \cap B^c) = P(A^c)\ .\ P(B^c)$$

الحـــل

من قانون ديمورجن $\quad A^cB^c = (A \cup B)^c$

$$P(A^cB^c) = P(A \cup B)^c = 1 - P(A \cup B)$$
$$P(A^cB^c) = 1 - [P(A) + P(B) - P(AB)]$$

ولكن A, B مستقلتان فإن :

$$P(A\,B) = P(A)\ .\ P(B)$$
$$\begin{aligned} P(A^cB^c) &= 1 - [P(A) + P(B) - P(A)\,.\,P(B)] \\ &= 1 - P(A) - P(B) + P(A)\,.\,P(B) \\ &= [1 - p(A)] - P(B) + P(A)\,.\,P(B) \\ &= [1 - P(A)] - P(B)\,[1 - P(A)] \\ &= P(A^c) - P(B)\,P(A^c) \\ &= P(A^c)\,[1 - P(B)] \\ P(A^cB^c) &= P(A^c)\ .\ P(B^c) \end{aligned}$$

وهو المطلوب .

مثال (٨ ـ ١٩)

أثبت أن :

$$P(A \cap B \cap C) = P(A)\,.\,P(B/A)\,.\,P(C/A \cap B)$$

الحـــل

$$\text{الطرف الأيمن} = P(A)\,.\,P(B/A)\,.\,P(C/A\cap B)$$

$$= P(A)\,.\,\frac{P(A\cap B)}{P(A)}\,.\,\frac{P(A\cap B\cap C)}{P(A\cap B)}$$

$$= P(A\cap B\cap C)$$

$$= \text{الطرف الأيسر}$$

مثال (٨ ـ ٢٠)

صندوق يحتوي على عشر كرات حمراء وعشرين كرة بيضاء أخذت عينة مكونة من كرتين واحدة بعد الأخرى أوجد احتمال أن تكون الكرتان لونهما أبيض.

ا) إن كان السحب بدون إرجاع.

ب) إن كان السحب بإرجاع.

الحـــل

$A = \{$ الكرة الأولى لونها أبيض $\}$

$B = \{$ الكرة الثانية لونها أبيض $\}$

$A\,B = \{$ الكرتان لونهما أبيض $\}$

ا) في حالة السحب بدون إرجاع

$$P(A\,B) = P(A)\,.\,P(B/A)$$

$$= \frac{20}{30}\,.\,\frac{19}{29}$$

ب) في حالة السحب بإرجاع

$$P(A\,B) = P(A)\,.\,P(B)$$

$$= \frac{20}{30}\,.\,\frac{20}{30}$$

مثال (٨ ـ ٢١)

حل المثال (٨ ـ ٢٠) إذا كانت العيّنة المسحوبة مكونة من ثلاث كرات.

الحــل

A = { الكرة الأولى بيضاء }

B = { الكرة الثانية بيضاء }

C = { الكرة الثالثة بيضاء }

ABC = { العيّنة كلها بيضاء }

ا) في حالة السحب بدون إرجاع

$$P(A\,B\,C) = P(A)\,P(B/A)\,P(C/AB)$$

$$= \frac{20}{30} \cdot \frac{19}{29} \cdot \frac{18}{28}$$

ب) في حالة السحب بإرجاع

$$P(A\,B\,B) = P(A) \,.\, P(B) \,.\, P(C)$$

$$= \frac{20}{30} \cdot \frac{20}{30} \cdot \frac{20}{30}$$

مثال (٨ ـ ٢٢)

في مثال (٨ ـ ٢٠) إذا سحبت عيّنة مكونة من 4 كرات بطريقة عشوائية.
ما هو احتمال الحصول على 3 كرات حمراء وواحدة لونها أبيض؟

الحــل

A = { العينة بها 3 كرات حمراء وواحدة بيضاء }

S = { العينة بها 4 كرات من 30 كرة }

باستخدام طرق العدّ

عدد الطرق للحصول على 3 كرات حمراء $= \binom{10}{3}$

عدد الطرق للحصول على كرة واحدة بيضاء $= \binom{20}{1}$

عدد الطرق للحصول على 4 كرات منها 3 حمراء، وواحدة بيضاء $= n(A)$

باستخدام قاعدة الضرب يكون :

$$n(A) = \binom{10}{3}\binom{20}{1}$$

عدد الطرق للحصول على 4 كرات من 30 كرة هي $n(S)$ حيث :

$$n(S) = \binom{30}{4}$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{\binom{10}{3}\binom{20}{1}}{\binom{30}{4}}$$

$$p(A) = \frac{10 \times 20 \times 24 \times 12}{30 \times 29 \times 28 \times 27} = 0.088$$

(٨ ـ ١٠) نظرية بيز (Bayes Theorem)

تهتم نظرية بيز بحساب احتمال أن يكون هناك سببًا ما هو مصدر حدوث حادثة معيّنة نعلم مسبقًا بحدوثها، وأن حدوثها يرجع إلى عدد من الأسباب المعروف احتمال حدوث كل منها، كما نعلم أيضًا احتمال حدوث الحادثة إذا تحقق سبب ما من هذه الأسباب.

إذا كان لدينا مجموعة الحوادث

$$A_1, A_2, \ldots, A_n$$

الشاملة والمتنافية بالتبادل (تجزيء فضاء العيّنة S) أي أن :

$$\bigcup_{i=1}^{n} A_i = S \;,\; A_i A_j = \phi \;,\; i \neq j = 1, \ldots, n$$

بحيث $P(A_i) \neq 0 \;,\; i = 1, \ldots, n$

لأي حادثة B معرفة على فضاء العيّنة S نفسه بحيث $P(B) \neq 0$ فإنه لقيم i التالية $i = 1, 2, \ldots, n$ يكون :

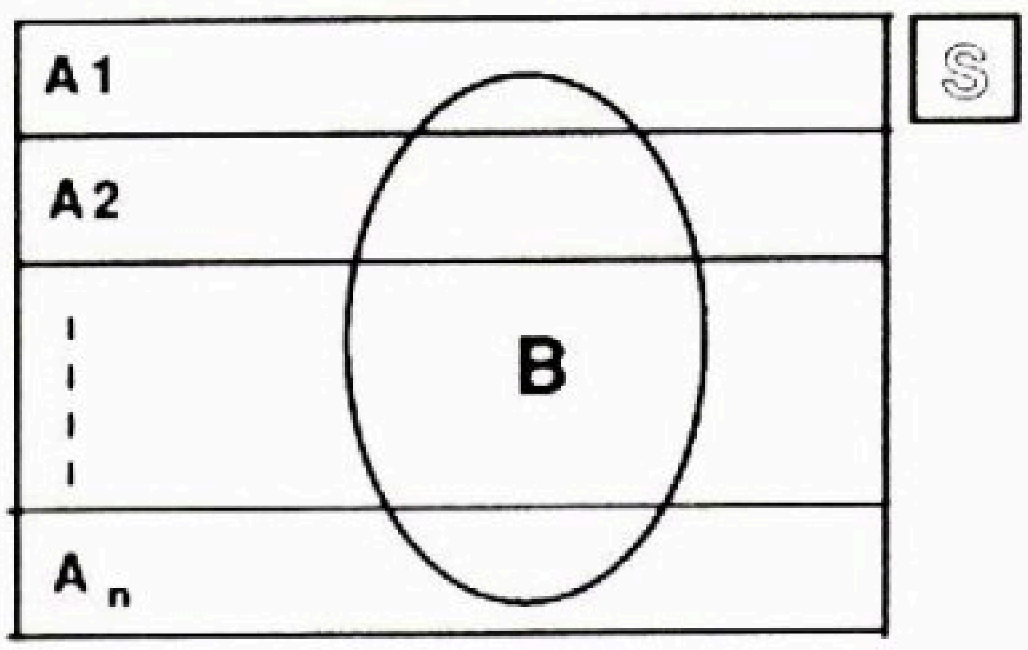
$$P(A_i / B) = \frac{P(B/A_i)\,P(A_i)}{P(B)} \;,\quad i = 1, 2, \ldots, n$$

أو

$$P(A_i / B) = \frac{P(B/A_i)P(A_i)}{\sum_{k=1}^{n} P(B / A_k) \, . \, P(A_k)} \qquad \text{.................} \qquad (٨ - ٦)$$

البرهان

يمكن توضيح الحوادث وفراغ العينة بشكل فن، شكل (٨ ـ ١١) التالي



شكل (٨ ـ ١١) يمثل شكل فن لتجزيء فراغ العينة S والحادثة B

من الشكل يتضح أن الحادثة B عبارة عن اتحاد مجموع الحوادث المتنافية التالية:

$$A_1B \, , A_2B \, , \ldots \, , A_nB$$

أي أن

$$B = A_1B \cup A_2B \cup \ldots \cup A_nB$$

وباستخدام المسلّمة الثالثة نحصل على:

$$P(B) = P(A_1B) + P(A_2B) + \ldots + P(A_nB)$$

$$= \sum_{k=1}^{n} P(A_kB)$$

$$= \sum_{k=1}^{n} P(B/A_k) \, P(A_k) \qquad \text{..............................} \qquad (1)$$

ومن تعريف الاحتمال الشرطي فإن

$$P(A_i / B) = \frac{P(A_iB)}{P(B)} = \frac{P(B/A_i)P(A_i)}{P(B)} \qquad \text{.......} \qquad (2)$$

من (2) ، (1) نحصل على :

$$P(A_i / B) = \frac{P(B/A_i)P(A_i)}{\sum_{k=1}^{n} P(B/A_k)P(A_k)} \quad ; \; i , k = 1 , \dots , n$$

وهو المطلوب .

ملاحظة مهمة

تسمى العلاقة

$$P(B) = P(B/A_1) P(A_1) + \dots + P(B/A_n) P(A_n)$$

$$\sum_{k=1}^{n} P(B/A_k) . P(A_k)$$

بقانون الاحتمال الكلي .

مثال (٨ ـ ٢٣)

مصنع به ثلاثة ماكينات I, II, III وكانت الماكينة I تنتج 20% من الإنتاج ، والماكينة II تنتج 30% من الإنتاج ، والماكينة III تنتج 50% من الإنتاج ، وكانت نسبة الإنتاج المعيب للماكينات الثلاث على الترتيب هو 4% و 3% و 2% .

فإذا اختيرت وحدة من الإنتاج بشكل عشوائي ، أحسب الاحتمالات التالية :

i) ما هو احتمال أن تكون الوحدة المسحوبة من الإنتاج معيبة؟

ii) إذا كانت الوحدة المسحوبة معيبة فما هو احتمال أن تكون من إنتاج الماكينة II .

الحل

نُعرّف الحوادث التالية

D = {الوحدة معيبة}

A = {الوحدة من إنتاج الماكينة I} ; $P(A) = 0.20$

B = {الوحدة من إنتاج الماكينة II} ; $P(B) = 0.30$

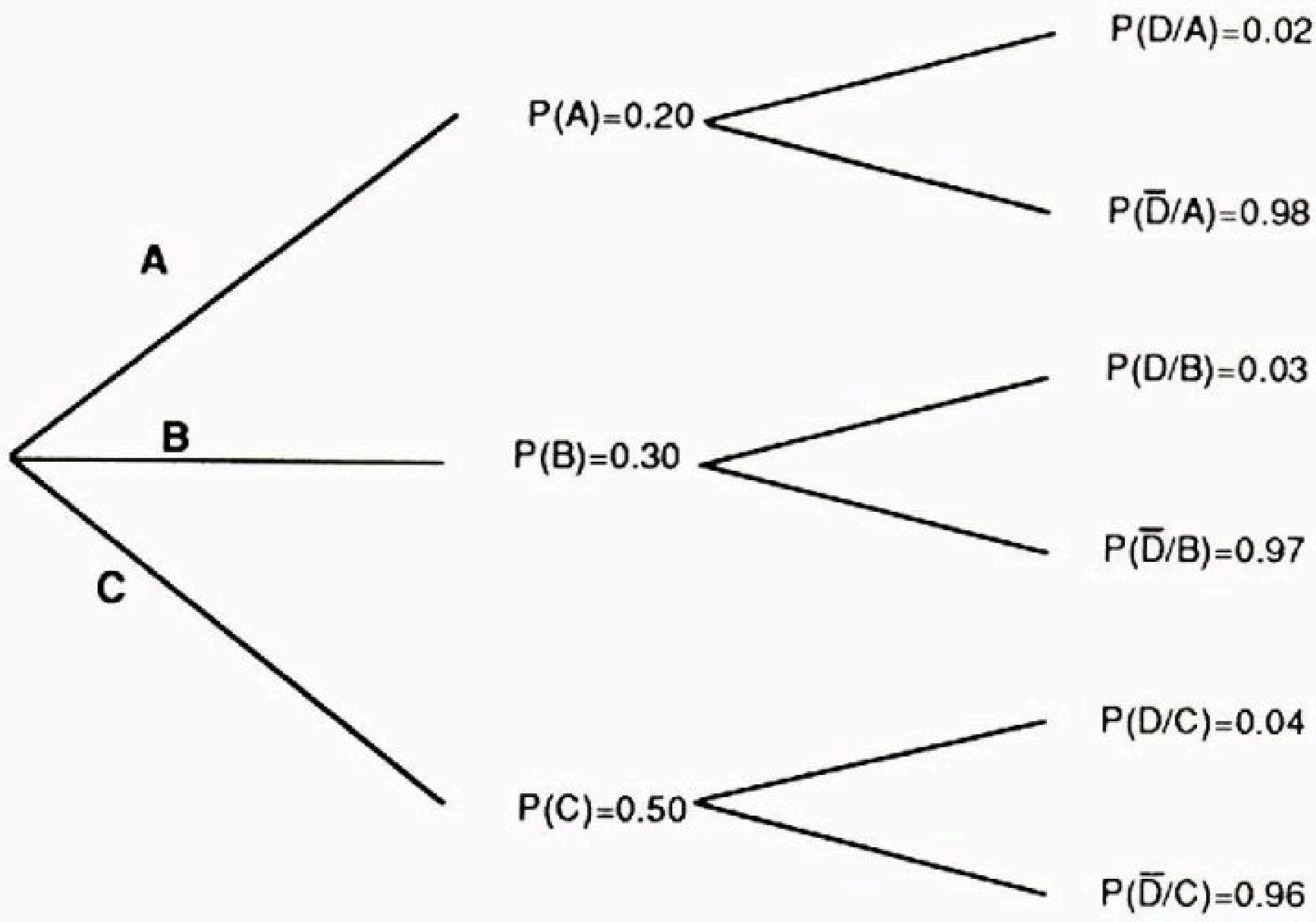
C = {الوحدة من إنتاج الماكينة III} ; $P(C) = 0.50$

$$P(D/A) = 0.02$$

$$P(D/B) = 0.03$$

$$P(D/C) = 0.04$$

ونوضح الحل باستخدام الشجرة البيانيةشكل (٨ ـ ١٢) التالي:



شكل (٨ ـ ١٢) يمثل الشجرة البيانية لاحتمالات الإنتاج للماكينات الثلاث

المطلوب هو إيجاد P(D)

(i

$$P(D) = P(A)P(D/A) + P(B)P(D/B) + P(C)P(D/C)$$

$$= 0.20 \times 0.02 + 0.30 \times 0.03 + 0.50 \times 0.04$$

$$= 0.004 + 0.009 + 0.020$$

$$= 0.033$$

(ii

$$P(B/D) = \frac{P(D/B)P(B)}{P(D)}$$

$$= \frac{0.03 \times 0.30}{0.033} = \frac{0.009}{0.033} = 0.273$$

مثال (٨ ـ ٢٤)

صندوقان الأول به 4 كرات بيضاء، 6 كرات سوداء والصندوق الثاني به 8 كرات بيضاء، 3 كرات سوداء. اختير أحد الصناديق عشوائيًا واختيرت منه كرة بطريقة عشوائية أوجد:

i) احتمال أن تكون الكرة المسحوبة لونها أسود.

ii) إذا اختيرت كرة ووجد أنها سوداء ما هو احتمال أن تكون من الصندوق الأول.

الحـــل

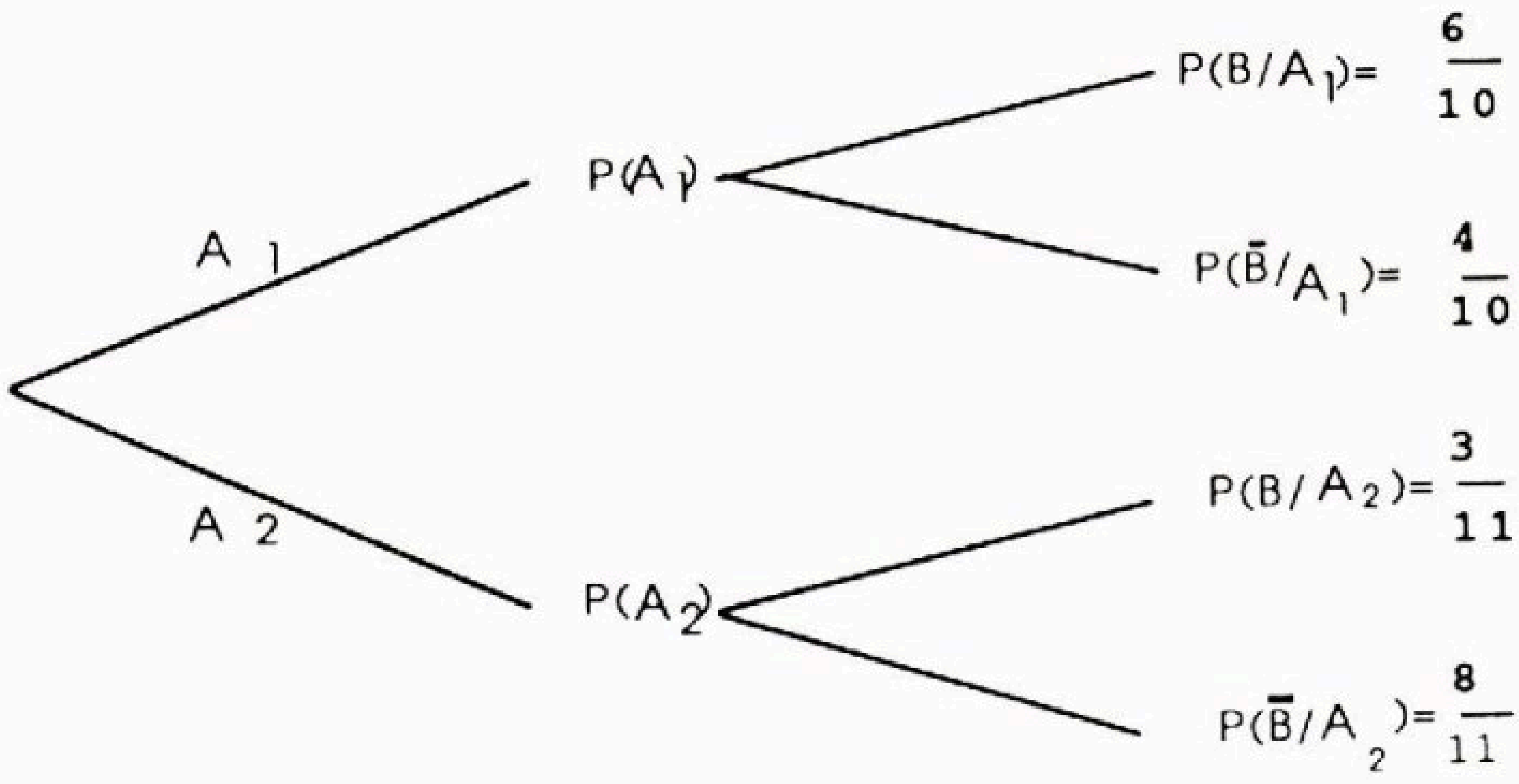
نفرض الحوادث التالية

A_1 = { الكرة المسحوبة من الصندوق الأول } , $P(A_1) = \frac{1}{2}$

A_2 = { الكرة المسحوبة من الصندوق الثاني } . $P(A_2) = \frac{1}{2}$

B = { الكرة المسحوبة سـوداء } . $P(B) = ?$

نوضح الحل باستخدام الشجرة البيانية كما شكل (٨ ـ ١٣) التالي:



شكل (٨ - ١٣) يمثل الشجرة البيانية لاحتمالات سحب كرة من أحد الصندوقين

i) المطلوب هو إيجاد $P(B)$

$$P(B) = P(B/A_1)\,P(A_1) + P(B/A_2)\,P(A_2)$$

$$P(B) = \frac{6}{10} \cdot \frac{1}{2} + \frac{3}{11} \cdot \frac{1}{2} = 0.436$$

ii) المطلوب هو $P(A_1/B)$

ويمكن إيجاد هذا الاحتمال باستخدام نظرية بيز كما يلي:

$$P(A_1/B) = \frac{P(B/A_1) \cdot P(A_1)}{P(B)}$$

$$= \frac{\frac{6}{10} \times \frac{1}{2}}{0.436} = 0.688$$

(٨ ـ ١١) تمـــارين

(١) أكتب عناصر فضاء العيّنات التالية :

i) $S = \{x :$ x عدد صحيح بين 1 ، 40 مقسومًا على 7 $\}$

ii) $S\{X : X^2 + X - 12 = 0\}$

iii) مجموعة النتائج الممكنة عند قذف حجر نرد وقطعة نقود على الترتيب .

iv) مجموعة النتائج الممكنة عند قذف حجري نرد مرة واحدة .

(٢) الحادثتان A, B معرفتان على قضاء العيّنة لتجربة عشوائية عبّر بكلمات واضحة للحوادث التالية :

$$A^c, B^c, A \cup B, A\ B, A^cB, AB^c, A^cB^c, A^c \cup B^c, (A \cup B)^c, (A\ B)^c$$

(٣) الحـوادث A, B, C معـرفة عل فضاء عيّنة لتجربة عشوائية . عبّر عن الحوادث التالية باستخدام الرموز

i) حدوث A مع عدم حدوث B مع عدم حدوث C .

ii) حدوث A, B مع عدم حدوث C .

iii) حدوث الحوادث الثلاث معًا .

iv) حدوث حادثة واحدة على الأقل .

v) حدوث حادثة واحدة فقط .

vi) عدم حدوث أي حادثة .

(٤) رمي حجر نرد مرة واحدة

i) أكتب فضاء العيّنة .

ii) أكتب الحادثة A الدالة على ظهور رقم أكبر من 3 .

iii) أكتب الحادثة B الدالة على ظهور رقم زوجي .

(٥) قذفت قطعتان نقديتان متزنتان معًا

i) أكتب فضاء العيّنة .

ii) أكتب الحوادث التالية :

A : الحادثة الدالة على ظهور صورة واحدة فقط .

B : الحادثة الدالة على ظهور صورة على الأقل .

C : الحادثة الدالة على ظهور كتابة في الرمية الأولى وصورة في الرمية الثانية .

(٦) رمي حجريْ نرد مرة واحدة

i) أكتب فضاء العيّنة .

ii) أكتب الحوادث التالية :

A الحادثة الدالة على ظهور رقمين متساويين .

B الحادثة الدالة على ظهور رقمين مجموعهما أكبر من 7 .

C الحادية الدالة على ظهور رقم 4 على الحجر الثاني .

iii) أكتب الحوادث التالية :

$A \cup B$, AB , AC , $A \cup C$

(٧) استخدم رسوم فن لإثبات العلاقات التالية :

i) $A \cup (AB) = A$

ii) $(AB) \cup (A\bar{B}) = A$

iii) $\overline{AB} = \bar{A} \cup \bar{B}$

iv) $A \cup B = (AB) \cup (A\bar{B}) \cup (\bar{A}B)$

v) $A(B \cup C) = (AB) \cup (AC)$

(٨) إذا كان الحادثتان A, B بحيث:

$$P(\bar{B}) = \frac{5}{8} \ , \ P(A \cup B) = \frac{3}{4} \ , \ P(A) = \frac{1}{2}$$

i) أحسب قيمة كل من:

$P(AB)$, $P(\bar{A} \cap \bar{B})$, $P(\bar{A} \cup \bar{B})$, $P(B\,\bar{A})$, $P(\bar{B}\,A)$, $P(A/B)$, $P(B/\bar{A})$

ii) هل A, B مستقلتان؟ أذكر السبب.

(٩) إذا كان $P(A) = 0.2 , P(A \cup B) = 0.7$

ا ـ أوجد P(B) في كل حالة من الحالات التالية:

i) A,B حادثتان مستقلتان.

ii) A,B حادثتان متنافيتان.

iii) الحادثة A حادثة جزئية من الحادثة B .

ب ـ إذا كان

$$P(C) = \frac{1}{3} , \ P(D/C) = \frac{1}{2} , \ P(C \cup D) = \frac{4}{5}$$

هل C,D مستقلتان؟ أذكر السبب.

(١٠) ا ـ بين الخطأ والصواب في كل مما يلي مع التعليل.

i) احتمال أن ينجح خالد في مقرر 101 إحص هو 0.7 واحتمال أن لا ينجح في هذا المقرر هو 0.2 .

ii) A,B حادثتان بحيث إن:

$$P(A) = 0.3 \ , \ P(B) = 0.4 \ , \ P(A \cap B) = 0.8$$

ب ـ A,B حادثتان بحيث إن:

$$P(B) = \frac{2}{3} , \ P(A \cap B) = \frac{1}{2} , \ P(A \cap B^c) = \frac{1}{4}$$

أحسب كلا من :

$$P(A)\ ,\ P(A\cap B)^c\ ,\ P(A^c\cap B^c)$$

(١١) مصنع لإنتاج مصابيح كهربائية سحبت منه عيّنة مكوّنة من ثلاث مصابيح واحد بعد الآخر فإذا رمزنا للمصباح المعيب بالرمز D وللمصباح السليم بالرمز G ، أكتب فضاء العيّنة وكذلك الحوادث التالية :

A = { العينة كلها معيبة }

B = { واحد على الأقل معيب }

C = { واحد على الأكثر معيب }

ثم أحسب :

$$A\cap B\ ,\ B\cup C\ ,\ \bar{B}\cap C$$

(١٢) صندوق به 15 قطعة من قطع غيار لنوع معيّن من الماكينات يحتوي على 10 قطع جيدة (G) و 5 قطع معيبة (D) . فإذا سحبنا عشوائيًا ثلاث قطع من الصندوق فما احتمال :

أ - أن تكون جميعها قطعًا جيدة .

ب - أن تكون جميعها معيبة .

جـ - أن تكون قطعتان جيدتين .

د - أن تكون على الأقل قطعتان جيدتين .

(١٣) وعاء به 4 كرات بيضاء و 6 كرات حمراء . سحبت كرة من الوعاء وأضيفت كرة من اللون المخالف للكرة المسحوبة ثم سحبت بعد ذلك كرة ثانية من الوعاء .

i) أوجد احتمال أن تكون الكرة الثانية حمراء .

ii) إذا كانت الكرتان المسحوبتان من اللون نفسه فما هو احتمال أن تكون الكرتان من اللون الأبيض .

(١٤) إذا علم أن احتمال أن يكون الجو ملبدًا بالغيوم هو 0.3 واحتمال أن يكون الجو إما ملبدًا بالغيوم أو عاصفًا هو 0.58 . أوجد الاحتمالات التالية إذا كان احتمال أن الجو عاصف هو 0.4 .

ا - أن يكون الجو ملبدًا بالغيوم وعاصفًا.

ب - أن يكون الجو ملبدًا بالغيوم وغير عاصف.

جـ - أن يكون الجو غير ملبد بالغيوم وغير عاصف.

(١٥) إذا علمت أن :

$$P(G) = 0.46 \ , \ P(H) = 0.53 \ , \ P(GH) = 0.21$$

أوجـد :

$$P(\bar{G}) \ , \ P(\bar{H}) \ , \ P(G \cup H) \ , \ P(\bar{G} \cap H) \ , \ P(\bar{G} \cup H) \ , \ P(\bar{G} \cap \bar{H})$$

(١٦) يحتوي صندوق على 12 علبة من الحليب المجفف 5 منها من الحجم الصغير و 4 منها من الحجم المتوسط و 3 منها من الحجم الكبير، واحدة من كل نوع خالية من الدسم والباقي كامل الدسم اخترنا علبة من الصندوق بصورة عشوائية ولنرمز بـ A لحادثة الحصول على علبة صغيرة و B لحادثة الحصول على علبة متوسطة و C لحادثة الحصول على علبة خالية من الدسم المطلوب.

ا - حساب الاحتمالات التالية :

$$P(A \cup B) \ , \ P(C) \ , \ P(C/A)$$

ب - هل الحادثتان A,C مستقلتان؟ ولماذا؟

(١٧) يحتوي صندوق على 9 قطع نقود من الأنواع ذات التواريخ المبينة في الجدول التالي :

ربع ريال	1376	1378	1400	1402
نصف ريال	1376	1400	1402	
ريـــال	1400	1403		

سحبنا قطعة بصورة عشوائية. لتكن A حادثة سحب ربع ريال، B حادثة سحب نصف ريال، C حادثة سحب قطعة نقود تحمل التاريخ 1400 والمطلوب حساب:

ا - $P(C)$, $P(C/A)$

ب - $P(C/B^{c})$

جـ - هل (A,C) مستقلتان؟ علل إجابتك.

د - هل (C,B) مستقلتان؟ علل إجابتك.

(١٨) صندوقان في الأول 4 كرات بيضاء و 6 كرات خضراء والثاني به 4 كرات بيضاء وكرة واحدة خضراء، أجب على الأسئلة التالية:

ا - سحبت كرة عشوائيًا من الصندوق الأول ما هو احتمال أن تكون خضراء.

ب - اختيرت عيّنة عشوائية من كرتين بدون إرجاع من الصندوق الأول، ما هو احتمال أن تكون من لونين مختلفين.

جـ - سحبنا عشوائيًا كرة من كل من الصندوق فما هو احتمال أن تكونا من اللون نفسه.

(١٩) بعد إجراء بعض الفحوصات الطبية على شخص وجد أن فشل الكلية اليسرى مستقل عن فشل الكلية اليمنى وكان احتمال فشل الكلية اليسرى 0.15 واحتمال فشل إحدى الكليتين على الأقل يساوي 0.2 .

أ - أوجد احتمال فشل الكلية اليسرى علمًا بأن الكلية اليمنى قد فشلت.

ب - أوجد احتمال عدم فشل الكلية اليمنى علمًا بأن الكلية اليسرى لم تفشل.

(٢٠) إذا كان لدينا حادثتان A,B بحيث كان:

$$P(A) = t \ , \ P(A \cup B) = \frac{1}{2} \ , \ P(B) = \frac{1}{3}$$

أوجد ما يأتي :

ا ـ قيمة t إذا كانت A,B متنافيتين .

ب ـ قيمة t إذا كانت A,B مستقلتين .

(٢١) إذا كان احتمال أن يصيب محمد هدفًا هو $\frac{1}{3}$ واحتمال أن يصيب أحمد الهدف نفسه هو $\frac{1}{5}$. أوجد احتمال أن يصيب واحد منهما على الأقل الهدف .

(٢٢) صنعت قطعة نقود بحيث إن احتمال ظهور الصورة $P(H) = \frac{1}{3}$ واحتمال ظهور الكتابة $P(T) = \frac{2}{3}$. ألقيت هذه القطعة مرة واحدة ثم نختار عددًا عشوائيًا من 1 إلى 11 إذا ظهرت صورة أما إذا ظهرت كتابة نختار بطريقة عشوائية عددًا من 1 إلى 7 . ما هو احتمال أن يكون العدد المختار فرديًا .

(٢٣) أثبت أن دالة الاحتمال الشرطي $P(./A)$ تحقق المسلمات الثلاثة لدالة الاحتمال .

(٢٤) وجد أن 0.4 من المراجعين في عيادة ما يشكون من ارتفاع في ضغط الدم وأن 0.2 من المراجعين مصابون بمرض في الكبد وأن 0.1 يشكون من المرضين معًا أوجد :

ا ـ احتمال أن أحد المراجعين يشكو من أحد المرضين على الأقل .

ب ـ هل ارتفاع ضغط الدم ومرض الكبد مستقلان؟

(٢٥) سحبت كرتان من صندوق به 15 كرة بيضاء ، 8 كرات سوداء ، سحبت عينة مكوّنة من كرتين واحدة بعد الأخرى بدون إرجاع ، أوجد الاحتمالات التالية :

ا ـ الكرتان من اللون نفسه .

ب ـ الكرتان لونهما أبيض .

جـ ـ الكرة الأولى بيضاء والثانية سوداء .

د ـ أوجد الاحتمالات السابقة إذا كان السحب بإرجاع .

(٢٦) تُرسل الإشارات اللاسلكية على شكل «نقاط» و «خطوط» حيث إن عدد النقاط = $\frac{3}{4}$ عدد الخـطوط. وبسبب الأخـطاء فإن النقـطة تصبح خطا باحتمال = $\frac{2}{3}$ والخط يصبح نقطة باحتمال $\frac{1}{4}$.

ا - ما احتمال استلام إشارة «نقطة»؟

ب - إذا استلمت إشارة «نقطة» فما احتمال أنها أرسلت «نقطة».

(٢٧) أعلنت إحدى الدوائر الحكومية عن حاجتها إلى عدد من الموظفين وبعد تصنيف مائة متقدم لهذه الوظيفة وفقًا للمؤهل ولسنوات الخبرة حصلنا على الجدول التالي :

المؤهل / الخبرة	يحمل شهادة جامعية	لا يحمل شهادة جامعية
ليس لديه خبرة	20	40
لديه خبرة	10	30

ا - اخترنا شخصًا بصورة عشوائية :

i) ما هو احتمال أن يكون ممن يحملون شهادة جامعية؟

ii) إذا علمت أن الشخص الذي اخترناه لم يكن لديه خبرة فما هو احتمال أن يكون من غير حملة الشهادة الجامعية.

ب - إذا اخترنا عشوائيًا شخصين على الترتيب، فما هو احتمال أن يكون الشخص الثاني ممن يحملون شهادة جامعية.

(٢٨) ثلاث ضاربي آلة كاتبة A, B, C يقومون بطباعة جميع الأوراق الخاصة بإحدى الشركات. الجدول التالي يبيّن النسب المئوية للأوراق التي يطبعها كل منهم والنسب المئوية للأخطاء التي يرتكبها كل منهم في عمله الخاص به.

الناسخون / النسب المئوية	النسبة المئوية للمطبوعات	النسبة المئوية للأخطاء
A	40 %	3 %
B	25 %	5 %
C	35 %	8 %

سحبنا ورقة بشكل عشوائي من مطبوعات الشركة :

ا - أوجد احتمال أن يوجد بها خطأ مطبعي .

ب - إذا علمت أن إحدى الأوراق المسحوبة تحوي خطأ طباعيًا فما احتمال أن B طبعها .

(٢٩) أربع طرق تؤدي إلى بئر ماء . اختار شخص أحد هذه الطرق عشوائيًا . فإذا اختار الطريق الأول A فإن احتمال وصوله إلى البئر يساوي $\frac{1}{8}$ وإذا اختار الطريق الثاني B فإن احتمال وصوله يساوي $\frac{1}{6}$. أما إذا اختار الطريق الثالث C إن احتمال وصوله يساوي $\frac{1}{4}$ وأخيرًا الطريق الرابع D فإن احتمال وصوله يساوي $\frac{9}{10}$.

والمطلوب :

ا - احتمال أن يصل الشخص إلى بئر الماء؟

ب - إذا نجح الشخص في الوصول إلى البئر، ما هو احتمال أن يكون قد اختار:

i) الطريق D ؟

ii) الطريق A ؟

(٣٠) ثلاث قطع نقود، الأولى متزنة والثانية والثالثة غير متزنتين . وإذا قذفنا الأولى فإن احتمال الحصول على صورة يساوي 0.5 ، وإذا قذفنا الثانية فإن احتمال

الحصول على صورة يساوي 0.4 ، وإذا قذفنا الثالثة فإن احتمال الحصول على صورة يساوي 0.75 . اخترنا إحدى القطع الثلاث عشوائيًا وقذفناها والمطلوب مايلي:

ا - أوجد احتمال الحصول على صورة.

ب - إذا علمت أن نتيجة القذفة كانت كتابة فما هو احتمال أن القطعة المقذوفة هي القطعة المتزنة.

(٣١) تنوي أسرة قضاء أجازة نهاية الأسبوع في إحدى الأماكن السياحية A أو B أو C . إذا كان احتمال سقوط المطر في A هو 0.6 وفي B هو 0.7 وفي C هو 0.5 وإذا اختارت الأسرة مكان الأجازة عشوائيًا فأحسب:

ا - احتمال أن تقضي الأسرة أجازة ممطرة.

ب - إذا علمت أن الأسرة قضت إجازة ممطرة فما هو احتمال أن أجازتها كانت في المكان B ؟

(٣٢) في إحدى مباني إسكان الجامعة يوجد 150 طالب منهم 111 يجيدون اللغة الإنجليزية، 50 يجيدون الفرنسية، 30 لا يجيدون أي لغة. اختير طالب عشوائيًا أوجد احتمال:

ا - أن يجيد اللغة الإنجليزية واللغة الفرنسية.

ب - أن يجيد اللغة الفرنسية علمًا بأنه لا يجيد اللغة الإنجليزية.

جـ - أن يجيد لغة واحدة على الأقل.

د - أن يجيد لغة واحدة فقط من الإنجليزية أو الفرنسية.

(٣٣) إذا كان نظام الدراسة في الدورة المكثفة للّغة الإنجليزية يقضي أن النتيجة إما راسب أو ناجح. نرمز للنجاح بالرمز S وللرسوب بالرمز F . اخترنا 3 طلاب منهم:

ا ـ عين فضاء العينة لنتائجهم . ثم عين نقاط الحوادث التالية :

{ أن ينجح اثنان منهم فقط } = A

{ ينجح واحد منهم على الأقل } = B

{ ألا ينجح أي منهم } = C

ب ـ إذا علمنا أن فرصة نجاح الطالب مثل فرصة رسوبه عين احتمالات الحوادث A, B, C .

جـ ـ أحسب احتمالات حدوث A علمًا بأن B قد وقعت . هل B, A مستقلتان . أذكر السبب .

(٣٤) ضع إشارة (✓) إذا كان الجواب صحيحًا أو إشارة (X) إذا الجواب خطأ :

ا ـ احتمال أن ينجح طالب في مقرر الإحصاء هو 0.9 واحتمال أن لا ينجح في هذا المقرر 0.15 ☐

ب ـ إذا كانت الحادثتان B، A معرفتين بحيث: $P(A) = 0.3, P(B) = 0.4.$ $P(A \cap B) = 0.2$ فإن :

(i) $P(A \cup B) = 0.7$ ☐

(ii) $P(\bar{A} \cap \bar{B}) = 0.5$ ☐

(iii) $P(\bar{A} / \bar{B}) = 0.65$ ☐

(iv) $P(\bar{A} \cup \bar{B}) = 0.8$ ☐

جـ ـ إذا كانت الحادثتان A، B مستقلتين فإن :

$P(A \cup B) = P(A) + P(B). P(\bar{A})$ ☐

(٣٥) أكمل كلا من الفراغات التالية :

ا ـ الحادثة هي ....

ب ـ إذا كان $P(A \cup B) = 0.85$ فإن $P(\bar{A} \cap \bar{B}) =$

جـ ـ إذا كان $P(A) = 0.3$ و $P(AB) = 0.2$ فإن $P(A\bar{B}) =$.

د - إذا سحبنا مع الإعادة كرتين من صندوق يتضمن خمس كرات بيضاء وخمس كرات حمراء فإن احتمال أن تكون الكرتان المسحوبتان من اللون الأبيض هو ....

هـ - الشرط اللازم والكافي لاستقلال الحادثتين A . B هو ....

و - الشرط اللازم والكافي لتنافي الحادثتين A . B هو ....

٩

المتغيرات العشوائية والتوزيعات الاحتمالية

Random Variable and Distribution

● مقدمة ● المتغير العشوائي المتقطع (المنفصل) ● دالة التوزيع الاحتمالي (دالـة الكتلة الاحتمالية) ● دالـة التـوزيـع الـتراكمي ● التـوقع للمتغير العشـوائي ● التباين للمتغير العشوائي ● بعض التـوزيعات المتقطعة الخاصة ● نظرية شيبشيف ● المتغير العشوائي المستمر (المتصل) والتوزيعات المتصلة ● التوزيع الطبيعي ● تمارين

(٩ - ١) مقدمـة

سبق لنا في الفصل السابق دراسة التجربة العشوائية وجميع النتائج الممكنة لها والتي تسعى فضاء العينة، وكذلك دراسة الحوادث التابعة لها وهي مجموعـات جزئيـة مـن فضاء العينة (الحادثة الأكيدة)، ودرسنا أيضًا احتمالات حدوث هذه الحـوادث. وفـي بعـض الأحيان قد لا يكون اهتمامنا بدراسة نقاط فضاء العينة ولكـن يكـون اهتمامنا منصبًّا على قيـم عدديـة مرتبطـة بنقاط فراغ العرينة، وهذه القيـم العدديـة المـرتبـطة بنقاط فضاء العينة هي ما تسمى بقيم المتغير العشوائي. وفي هذا الفصل سنـدرس المتغـيرات العشـوائية، وكذلك التوزيعات الاحتمالية والتوقع والتباين لهذه المتغيرات العشوائية كما يلي.

(٩ ـ ٢) المتغيّر العشوائي (Random variable)

هو دالـة ذات قيم عددية حقيقية معرفة على فضاء العيّنة S . أي أن المتغيّر العشوائي هو تطبيق مجاله فضاء العيّنة S ، وقيمـه أو مجالـه المقابل X(S) هو مجموعة جزئية من مجموعة الأعداد الحقيقية R .

ملاحظـات

١ ـ تعرف الصورة العكسية للنقطة $x \in R$ تحت تأثير المتغير العشوائي X ويرمـز لهـا بالرمز A_x على النحو التالي:

$$A_x = \{w \in S : X(w) = x\}$$

لذلك فإن الصورة العكسية للمجموعة $X(s) \subset R$ هي $\bigcup_{x \in X(s)} A_x$.

٢ ـ يعرف حقل بوريل (ويرمز له بالرمز $\mathscr{B}$) على أنه أصغر حقل سيجما ξ يحتوي المجموعات على الصورة $\{x : a < x < b\}$ $a, b \in R$

٣ ـ هناك تعريف متقدم للمتغيّر العشوائي كالتالي:
يقال إن X متغير عشوائي على S إذا كانت جميع الصور العكسية لكل عنصر في $\mathscr{B}$. موجودة في حقل سيجما ξ المعرّف على (S) .

٤ ـ في هـذا الكتـاب سنكتفي باعتبـار مجموعة القوى P_s للمجموعة S هـي حقـل سيجما ξ المعرّف على فضاء العيّنة S .

والمثال التالي يعرف أحد المتغيرات العشوائية الممكنة على فضاء عينة يعطى ثـم يبـينّ كيفيـة استخراج الصور العكسية ويوضح أن الصور العكسية تكون تجزيئًا لفضـاء العيّنـة ثـم يوضّـح طرق حساب بعض قيم الاحتمالات الخاصة بالمتغيّر العشوائي .

مثال (٩ ـ ١)

اعتبر تجربة قذف قطعة عملة متّزنة مرتين واعتبر المتغيّر العشوائي X هو عدد الصور الناتجة من القذف. عبّر عن هذا المتغيّر العشوائي كتطبيق مجاله فضاء العيّنة ثم بيّن الصور العكسية لكل قيمة عددية لهذا المتغيّر ووضح أن هذه الصور تشكل تجزيء لفضاء العيّنة واحسب $P(X<2)$.

الحـــل

نكتب فراغ العيّنة S كالتالي:

$$S = \{(H,H), (H,T), (T,H), (T,T)\}$$

أو نكتب بشكل آخر

$$S = \{w_1, w_2, w_3, w_4\}$$

حيث

$$w_1 = (H,H) \;,\; w_2 = (H,T) \;,\; w_3 = (T,H) \;,\; W_4 = (T,T)$$

والاحتمالات

$$P(\{w_1\}) = P(\{w_2\}) = P(\{w_3\}) = P(\{w_4\}) = \frac{1}{4}$$

أو بشكل آخر

$$P_1 = P_2 = P_3 = P_4 = \frac{1}{4}$$

فإذا عرّفنا المتغير X = عدد الصور، فيعبّر عن هذا المتغيّر كتطبيق مجاله S كالآتي:

$$X(w_1) = X(H,H) = 2$$

$$X(w_2) = X(H,T) = 1$$

$$X(w_3) = X(T,H) = 1$$

$$X(w_4) = X(T,T) = 0$$

وتكون الصور العكسية كالآتي:

$$A_0 = \{w_4\} \;,\; A_1 = \{w_2, w_3\} \;,\; A_2 = \{w_1\}$$

هي مجموعة جزئية من S بحيث

$$\text{(i)}\quad A_i \cap A_j = \phi \quad , \quad i \neq j \qquad i,j = 0,1,2$$

(ii) $A_0 \cup A_1 \cup A_2 = S$

(iii) A_0, A_1, A_2 تسمى فضاء الحدث للمتغيّر العشوائي X

ونلاحظ من (i), (ii), (iii) أن A_0, A_1, A_2 تشكّل تجزيئًا لفضاء العيّنة S والمجموعة $X(s)$ والتي هي مجموعة جزئية من R .

$$X(s) = \{\{0, 1, 2\} \subset R$$

تسمى مجموعة القيم الممكنة للمتغيّر العشوائي X (أو المجال المقابل $X(s)$) .
المصطلح $(X = x)$ سوف يستخدم كاختصار للحادثة A_x أي أن :

$$(X = x) = \{w \in S / X(w) = x\}$$

أو بشكل عام

$$(X < x) = \{w \in S / X(w) < x\}$$

$$(X \text{ in } E) = [w \in S / X(w) \text{ in } E]$$

وحيث إن $(X = x)$ [أو بشكل عام $(X \text{ in } E)$] هي حادثة فإن $P(X = x)$ [أو بشكل عام $P(X in E)$] تكون معرّفة بشكل جيد وهو احتمال مجموعة نقاط العينة w والتي تحقق $X(w) = x$ [أو بشكل عام $(X(w) \text{ in } E)$] وعودة إلى المثال (٩ ـ ١) السابق نجد أن :

$$(X = 2) = \{w_1\}$$

$$(X < 2) = \{w_2, w_3, w_4\}$$

$$P(X < 2) = P(\{w_2, w_3, w_4\}) = P(\{w_2\} \cup \{w_3\} \cup \{w_4\}$$

$$= P(\{w_2\}) + P(\{w_3\}) + P(\{w_4\})$$

$$= P_2 + P_3 + P_4 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

(٩ ـ ٢ ـ ١) المتغيّر العشوائي المتقطع أو المنفصل (discrete random variable)

هو المتغيّر الذي تكون مجموعة قيمه الممكنة هي مجموعة متقطّعة أو بعبارة أخرى إذا كان مجاله المقابل $X(s)$ قابل للعدّ (قيمه الممكنة قابلة للعدّ) .

ومن أمثلة المتغيّر العشوائي المتقطّع، عدد الصور التي تظهر عند قذف قطعة نقود n من المرات، قيمة الرقم الذي يظهر عند رمي حجر النرد، عدد الوحدات المعيبة عند سحب عينة n من إنتاج إحدى الآلات، عدد السفن التي تصل أحد الموانىء خلال أسبوع، عدد حوادث المرور في أحد الشوارع في مدينة ما خلال شهر معيّن. حجم الأسرة عند سحب عيّنة من الأسر بإحدى المدن وهكذا. وسوف نرمز للمتغيّرات العشوائية بالحروف الكبيرة ... X, Y, Z, وقيم هذه المتغيّرات بالحروف الصغيرة ونوضح ذلك بالأمثلة التالية.

مثال (٩ ـ ٢)

قذفت قطعة عملة مرتين متتاليتين. أوجد القيم الممكنة للمتغيرات العشوائية التالية:

(١) X = عدد الصور.

(٢) Y = مربع عدد الصور.

(٣) Z = الفرق بين عدد ظهور الصورة وعدد ظهور الكتابة.

الحـــل

فضاء العيّنة S لقذف قطعة العملة مرتين هو:

$$S = \{HH, HT, TH, TT\}$$

أو

$$S = \{w_1, w_2, w_3, w_4\}$$

المتغيّر العشوائي X تحسب قيمه الممكنة كالتالي:

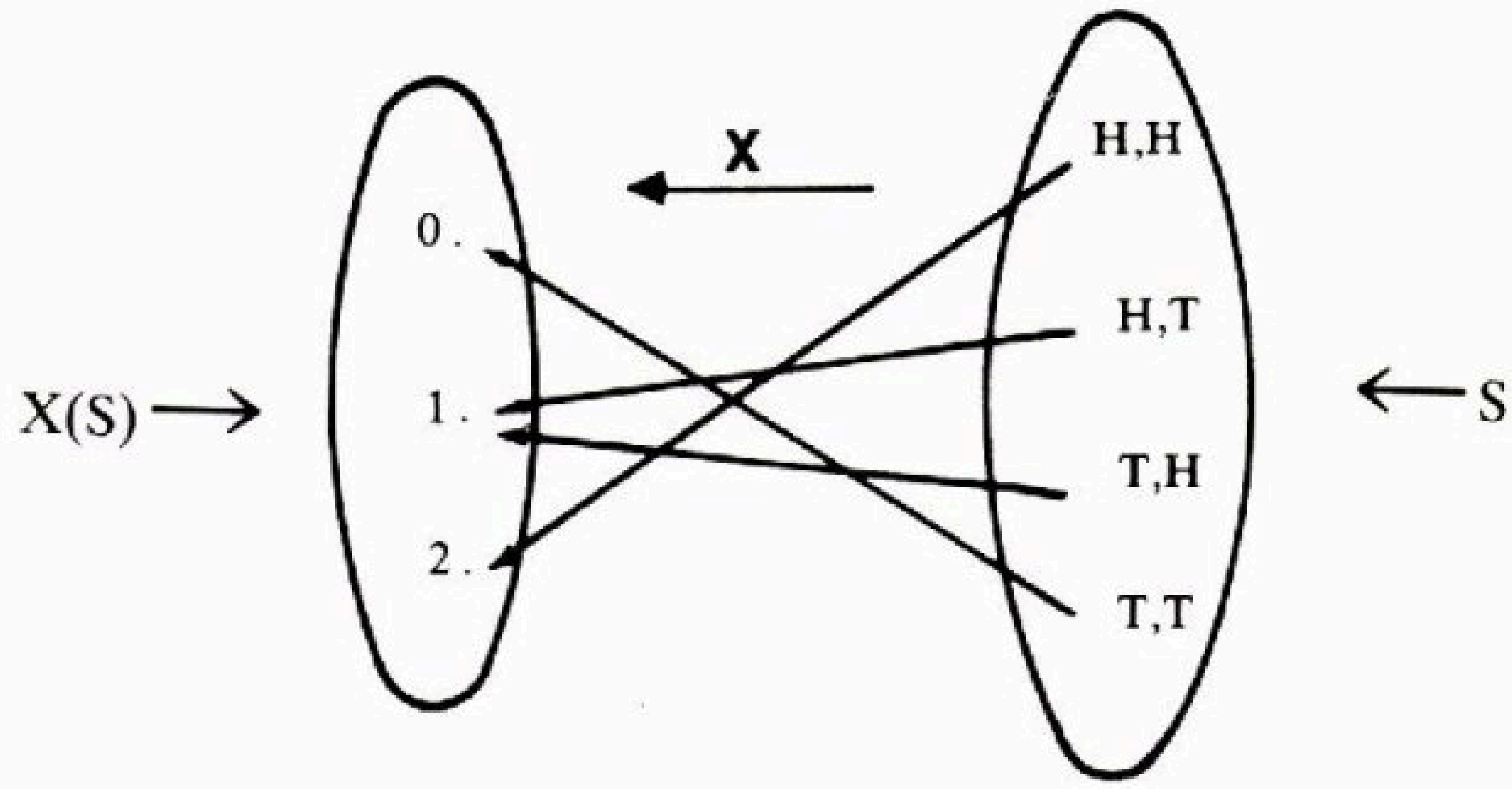
$$X(w_1) = X(H,H) = 2$$

$$X(w_2) = X(H,T) = 1$$

$$X(w_3) = X(T,H) = 1$$

$$X(w_4) = X(T,T) = 0$$

وعليه فإن القيم الممكنة للمتغيّر العشوائي X هي $x = 0, 1, 2$ ونوضح ذلك بالرسم، شكل (٩ ـ ١) كالتالي:



شكل (٩ ـ ١) يمثل المجال والمجال المقابل للمتغيّر العشوائي X

المتغيّر العشوائي Y نحسب قيمه الممكنة كالتالي :

$$Y(w_1) = Y(H,H) = 2^2 = 4$$

$$Y(w_2) = Y(H,T) = 1^2 = 1$$

$$Y(w_3) = Y(T,H) = 1^2 = 1$$

$$Y(w_4) = Y(T,T) = (0)^2 = 0$$

وعليه فإن القيم الممكنة للمتغيّر العشوائي Y هي : $y = 0, 1, 4$

المتغيّر العشوائي Z نحسب قيمه الممكنة كالتالي :

$$Z(w_1) = Z(H,H) = 2 - 0 = 2$$

$$Z(w_2) = Z(H,T) = 1 - 1 = 0$$

$$Z(w_3) = Z(T,H) = 1 - 1 = 0$$

$$Z(w_4) = Z(T,T) = 0 - 2 = -2$$

وعليه فإن القيم الممكنة للمتغيّر العشوائي Z هي $Z = \{-2, 0, 2\}$

مثال (٩ ـ ٣)

أحسب من مثال (٩ ـ ١) قيمة الاحتمال $P(X<2)$ في الحالتين التاليتين :

الحالة الأولى : إذا كانت قطعة العملة متزنة [أي أن : $P(H) = P(T) = \frac{1}{2}$]

الحالة الثانية : إذا كانت قطعة العملة غير متزنة عندما $P(H) = \frac{1}{3}$, $P(T) = \frac{2}{3}$

الحـــل

الحالة الأولى:

$$P_1 = P_2 = P_3 = P_4 = P(H,H) = P(H)\,P(H) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$P(X < 2) = P_2 + P_3 + P_4 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

الحالة الثانية:

$$P_1 = P(w_1) = P(H,H) = P(H)\,P(H) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

$$P_2 = P(w_2) = P(H,T) = P(H)\,P(T) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

$$P_3 = P(w_3) = P(T,H) = P(T)\,P(H) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$$

$$P_4 = P(w_4) = P(T,T) = P(T)\,P(T) = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$$

$$P(X < 2) = P_2 + P_3 + P_4$$

$$= \frac{2}{9} + \frac{2}{9} + \frac{4}{9} = \frac{8}{9}$$

(٩ ـ ٣) دالة التوزيع الاحتمالي (دالة الكتلة الاحتمالية)

Probability mass function

إذا كان X متغيّر عشوائي متقطّع (غير مستمر) ويأخذ القيم القابلة للعدّ التالية:

$$x_1, x_2, \ldots$$

فالدالة $f_x(.)$ والمُعرّفة بالقانون التالي:

$$f_X(x_i) = \begin{cases} P(X = x_i) & , X = x_i \; ; \; i = 1, 2, \ldots \\ 0 & , X \neq x_i \end{cases} \qquad (٩ - ١)$$

تسمى دالـة كتلة احتـمالية للمتغـيّر العشوائي X أو دالة التكرار المتقطعة إذا كان $\sum_i f_x(x_i) = 1$

ويمكن أن نوضح قيم المتغيّر العشوائي المتقطّع X ودالة توزيع الكتلة $f_x(.)$ له في الجدول التالي:

x	x_1	x_2	....	x_n
$f(x)$	$f(x_1)$	$f(x_2)$	....	$f(x_n)$

مثال (٩ ـ ٤)

أوجد دالة توزيع الكتلة الاحتمالية $f_X(.)$ للمتغيّر العشوائي X في مثال (٩ ـ ١) ثم وضّح الإجابة بالرسم البياني (في حالة قطعة العملة المتزنة).

الحـــل

وحيث إن قيم المتغــير العشـــوائي X هي $\{0, 1, 2\}$ فإن دالـــة الكتلة الاحتـمالية $f_X(.)$ تحسب كما يلي:

$$f_X(0) = P(X = 0) = P\{[TT\}) = \frac{1}{4}$$

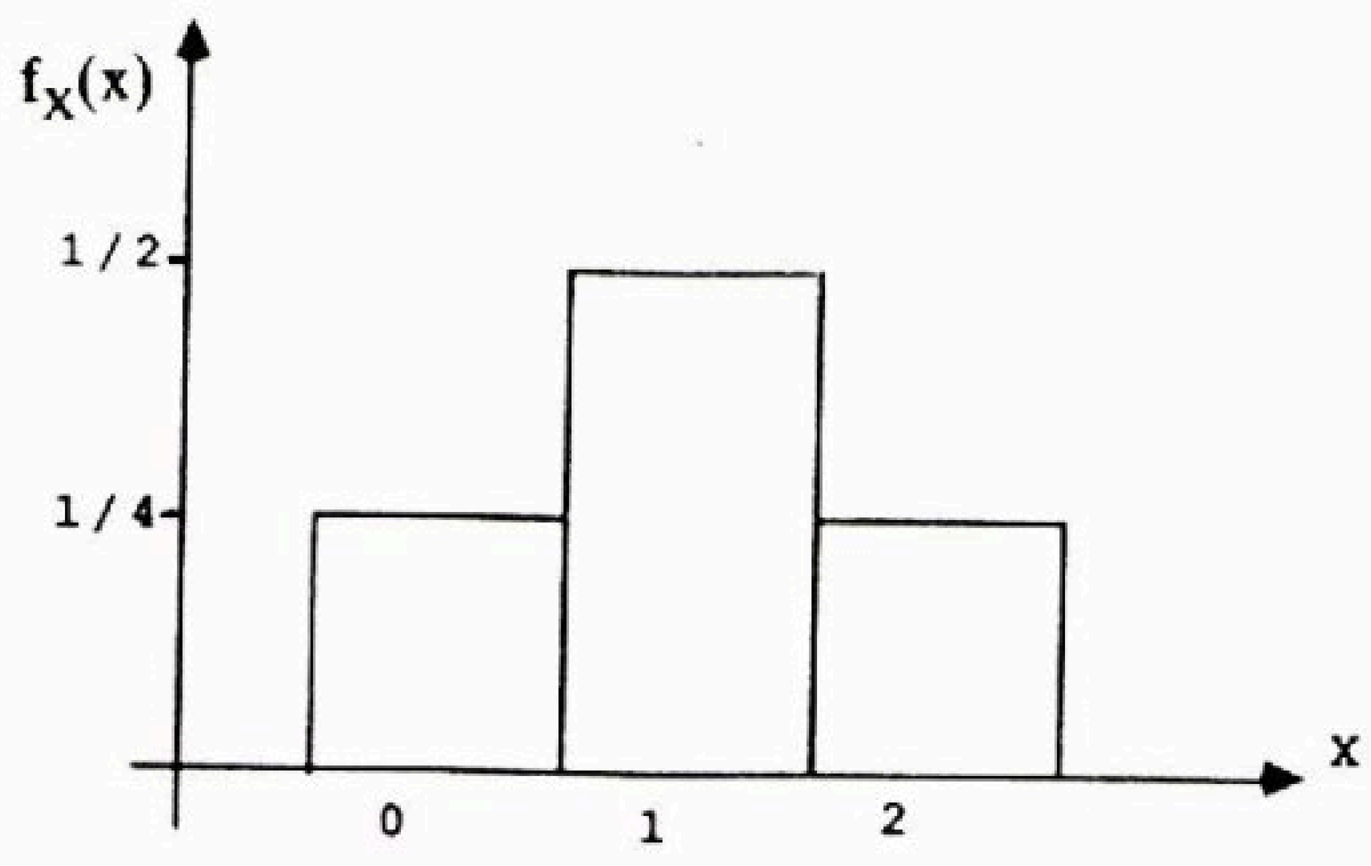
$$f_X(1) = P(X = 1) = P\{[HT, TH\}) = \frac{2}{4} = \frac{1}{2}$$

$$f_X(2) = P(X = 2) = P(\{HH\}) = \frac{1}{4}$$

ويمكن وضـع قيم المتغيّر العشوائي X وقيم دالة الكتلة الاحتمالية المناظرة في الجدول التالي:

x	0	1	2
$f_X(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

ويمكن توضيح الجدول بيانيًا في شكل (٩ـ٢) التالي:



شكل (٩ ـ ٢) يمثل المدرج التكراري لدالة الكتلة الاحتمالية $f_X(x)$

(٩ ـ ٣ ـ ١) خواص دالة الكتلة الاحتمالية $f_X(x)$

دالة توزيع الكتلة $f_X(x)$ للمتغيّر العشوائي X يكون لها الخاصيتان التاليتان:

ا) جميع قيم دالة توزيع الكتلة $f_X(x)$ تكون غير سالبة أي أن $f_X(x) \geqslant 0$.

ب) مجمـوع قيم دالة توزيع الكتلة $f_X(x)$ المصاحبة لجميع قيم المتغيّر العشوائي X تكون مساوية للواحد الصحيح أي أن:

$$\sum_{\text{all } x} f_X(x) = 1$$

مثال (٩ ـ ٥)

في مثال (٩ ـ ١) أثبت أن خاصيتي دالة توزيع الكتلة السابقة محققة وذلك في حالة قطعة العملة المتّزنة.

الحـــل

من الجدول في مثال (٩ ـ ٤) نلاحظ أن:

$$f_X(0) = \frac{1}{4} > 0 \quad , \quad f_X(1) = \frac{1}{2} > 0 \quad , \quad f_X(2) = \frac{1}{4} > 0$$

أي أن الخاصية (ا) محققة.

وأن

$$\sum_{i=1}^{3} f_X(x_i) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

أي أن الخاصية (ب) محققة أيضًا.

(٩ ـ ٤) دالة التوزيع التراكمي
Cumulative distribution function

إذا كان لدينا متغيّر عشوائي يأخذ القيم $x_1, x_2, \ldots, x_n, \ldots$ ودالة كتلته الاحتمالية المناظرة هي $f_X(x_1), f_X(x_2), \ldots, f_X(x_n), \ldots$ فإن دالة التوزيع التراكمي $F_X(.)$ للمتغيّر X هي:

$$f_X(x_i) = \sum_{x_i \leq x} f_X(x_i) \quad ; \quad -\infty < x < \infty \qquad \ldots\ldots\ldots\ldots \quad (٩ - ٢)$$

(٩ ـ ٤ ـ ١) خصائص دالة التوزيع التراكمي

دالة التوزيع التراكمي يكون لها الخواص التالية:

١) $F_X(x) = P(X \leq x)$

وينتج من هذا

$$F_X(-\infty) = 0 \quad ; \quad F(+\infty) = 1$$

الإثبات

$$F_X(x) = \sum_{x_i \leq x} f(x_i) = P(X \leq x)$$

٢) دالة التوزيع التراكمي متزايدة (غير متناقصة) أي أن

$$F_X(a) \leq F_X(b) \quad ; \quad a < b$$

الإثبات

نفرض $a < b$

الحدث $\{X \leq b\}$ يمكن كتابته كاتحاد حدثين متنافيين كالتالي:

$$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\}$$

وحيث إن $\{X \leq a\} \cap \{a < X \leq B\} = \phi$

$$P(X \leq b) = P(X \leq a) + (P(a < X \leq b)$$

أي

$$P(X \leq b) \geq P(X \leq a)$$

لأن

$$P(a < X \leq b) \geq 0$$

وعليه فإن

$$F(b) \geq F(a)$$

وهو المطلوب.

٣)

$$0 \leq F(x) \leq 1$$

من الخاصية (٢) نجد أن

$$F(-\infty) \leq F(x) \leq F(\infty)$$

أي أن:

$$0 \leq F(x) \leq 1$$

مثال (٩ ـ ٦)

من مثال (٩ ـ ١) السابق أحسب دالة التوزيع التراكمي $F(x)$ لقيم x المختلفة ثم وضح ذلك بالرسم البياني:

الحـــل

من التعريف (٩ ـ ١)

$$F_X(x) = \sum_{x_i \leq x} f_X(x_i), \qquad -\infty < x < \infty$$

نجد أن

$$F_X(-\infty) = 0$$

وذلك من الخاصية (ا)

لأي $-\infty < x < 0$ فإن $F_X(x) = 0$

لأي $0 \leq x < 1$ فإن $F_X(x) = f_X(0) = \frac{1}{4}$

لأي $1 \leq x < 2$ $F_X(x) = f_X(0) + f_X(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$,

لأي $2 \leq x < \infty$ $F_X(x) = f_X(0) + f_X(1) + f_X(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$,

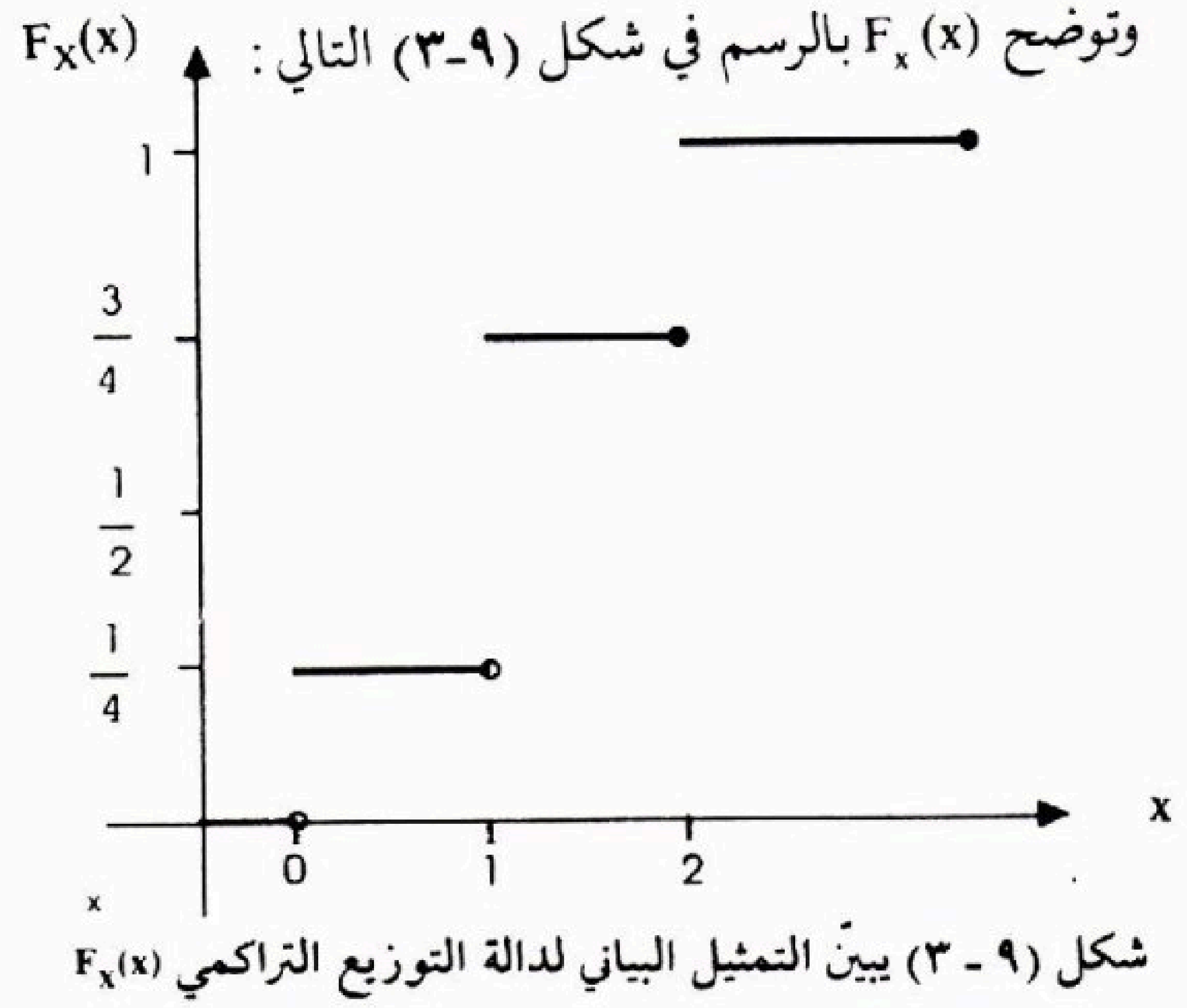
ومن الخاصية (ا)

$$F_X(\infty) = 1$$

ويلخص هذا في شكل دالة كالتالي :

$$F_X(x) = \begin{cases} 0 & -\infty \leq x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

وتوضح $F_x(x)$ بالرسم في شكل (٩-٣) التالي :



شكل (٩ - ٣) يبيّن التمثيل البياني لدالة التوزيع التراكمي $F_X(x)$

(٩ ـ ٥) التوقع للمتغيّر العشوائي

إذا كان لدينا المتغيّر المتقطع X يأخذ القيم المنتهية التالية:

$$x_1, x_2, \ldots, x_n$$

ودالة الكتلة الاحتمالية المناظرة هي:

$$f(x_1), f(x_2), \ldots, f(x_n)$$

فإن التوقع للمتغيّر العشوائي X أو المتوسط (القيمة المتوقعة) له، التي ترمز لها بالرمز $E(X)$ أو μ_X أو μ تعطى كما يلي:

$$\mu_X = E(X) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_n f(x_n)$$

أو باختصار كالتالي:

$$\mu_X = E(X) = \sum_{i=1}^{n} x_i f(x_i) \qquad \text{(٩ ـ ٣)}$$

مثال (٩ ـ ٧)

أحسب التوقع $E(X)$ للمتغيّر العشوائي X في مثال (٩ ـ ١).

الحـــل

$$\begin{aligned} E(X) &= \sum_{x=0}^{2} x\, f_X(x) \\ &= (0) \,.\, f_X(0) + (1) \,.\, f_X(1) + (2) \,.\, f_X(2) \\ &= 0 + 1\left(\frac{1}{2}\right) + 2\left(\frac{1}{4}\right) = 1 \end{aligned}$$

(٩ ـ ٥ ـ ١) بعض خواص التوقع

١) توقع المقدار $aX \pm b$ حيث b, a ثوابت يعطى كما يلي:

$$E(aX \pm b) = aE(X) \pm b$$

لأن

$$\begin{aligned} E(aX \pm b) &= \sum_{x} (ax \pm b) f_X(x) \\ &= \sum_{x} ax f_X(x) \pm \sum_{x} b f_X(x) \end{aligned}$$

$$= a\sum_{x} xf_X(x) \pm b\sum_{x} f_X(x)$$

$$= aE(X) \pm b$$

حيث $\sum_{x} f_X(x) = 1$ من الخاصية (ب) لدالة توزيع الكتلة.

ب) توقع المقدار الثابت b يساوي المقدار الثابت نفسه أي أن:

$$E(b) = b$$

لأن

$$E(b) = \sum_{x} bf_X(x) = b\sum_{\chi} f_X(x) = b$$

جـ) توقع المتغيّر العشوائي X مضروبًا في مقدار ثابت يساوي المقدار الثابت مضروبًا في توقع المتغيّر العشوائي أي أن:

$$E(a\,X) = aE(X)$$

ويمكن الحصول على هذه الخاصية بجعل b تساوي صفرًا في الخاصية (ا).

د) إذا كان $f_X(x)$ دالة الكتلة الاحتمالية للمتغيّر العشوائي X وكانت الدالة G(X) دالة في المتغيّر العشوائي X فإن:

$$E\big(G(X)\big) = \sum_{x} G(x)\, f_X(x)$$

(٩ ـ ٦) التباين للمتغيّر العشوائي

تباين المتغيّر العشوائي X الذي له متوسط μ_X ويرمز له بالرمز V(X) أو σ_X^2 ويعطى بالعلاقة:

$$\sigma_X^2 = E\,[(X-\mu_X)^2] \qquad \text{.................. (٩ ـ ٤)}$$

والجذر التربيعي للتباين يسمى بالانحراف المعياري ونرمز له بالرمز σ_X ويحسب التباين V(X) للمتغيّر العشوائي X الذي له القيم

$$x_1, x_2, \ldots, x_n, \ldots$$

ودالة توزيع الكتلة المناظرة لهذه القيم على الترتيب هي :

$$f_X(x_1)\ , f_X(x_2), \ldots , f_X(x_n), \ldots$$

فإن التباين يعطى كما يلي :

$$V(X) = E(X-\mu_X)^2 = \sum_x (x-\mu_X)^2 f_X(x)$$

وباستخدام خواص التوقع يمكن أن يكتب $V(X)$ كالتالي :

$$V(X) = E(X^2) - \mu^2$$

حيث

$$E(X^2) = \sum_x x^2 f(x)$$

لأن

$$V(X) = E[(X-\mu_X)^2] = \sum_x (x-\mu_X)^2 f(x)$$

$$= \sum_x (x^2 - 2x\mu_X + \mu_X^2) f(x)$$

$$= \sum_x x^2 f(x) - 2\mu_X \cdot \sum_x x f(x) + \mu_X^2 \Sigma f(x)$$

$$= \sum_x x^2 f(x) - \mu_X^2 \qquad \text{(٩ - ٥)}$$

ملاحظة

إذا فسّرنا التوزيع التكراري بتوزيع الكتلة بالنسبة لجسم كتلته الوحدة، واعتبرنا أن الكتلة $f_X(x_i)$ موضوعة عند النقطة $x = x_i$ فإن :

١) القيمة المتوقعة $\mu = E(X)$ يناظر مركز ثقل هذا الجسم $\bar{x}$ حيث

$$\bar{x} = \frac{\Sigma x_i f(x_i)}{\Sigma f(x_i)} = \Sigma x_i f(x_i) = E(X)$$

٢) التباين لـ X يناظر القصور الذاتي I حيث تباين في هذه التكرارات هو:

$$I = \frac{\Sigma (x_i - \mu)^2 f(x_i)}{\Sigma f(x_i)} = \Sigma (x_i - \mu)^2 f(x_i) = \sigma_X^2$$

مثال (٩ ـ ٨)

أحسب التباين والانحراف المعياري للمتغيّر العشوائي X في مثال (٩ ـ ١).

الحـــل

$$V(X) = E(X^2) - \mu^2$$

في مثال (٩ ـ ٧) حسبنا قيمة μ_X وتساوي 1 والآن نحسب $E(X^2)$

$$\begin{aligned} E(X^2) &= \sum_{x} x^2 f_X(x) \\ &= (0)^2 . f_X(0) + (1)^2 . f_X(1) + (2)^2 . f_X(2) \\ &= 1. (\frac{1}{2}) + 4. (\frac{1}{4}) = 1.5 \end{aligned}$$

$$V(X) = 1.5 - (1)^2 = 0.5$$

$$\sigma_X = \sqrt{0.5} = 0.7$$

(٩ ـ ٦ ـ ١) بعض خواص التباين

ا) تباين المقدار الثابت يساوي صفرًا أي أن:

$$V(C) = 0$$

لأن

$$V(C) = E(C - \mu_c)^2 = E(C - C)^2 = E(0) = 0$$

ب) تباين المقدار $aX \pm b$ حيث a,b مقداران ثابتان، X متغيّر عشوائي يساوي مربع المقدار a مضروبًا في تباين المتغير X أي أن:

$$V(aX \pm b) = a^2 V(X)$$

لأن

$$\begin{aligned} V(aX \pm b) &= E(aX \pm b - a\mu_X \mp b)^2 \\ &= a^2 E(X - \mu_X)^2 = a^2 V(X) \end{aligned}$$

(٩ ـ ٧) بعض التوزيعات المتقطّعة الخاصّة

سبق لنا دراسة المتغيّر العشوائي المتقطّع ودالة توزيع الكتلة والتوقع والتباين بصورة عامة والآن سوف ندرس بعض التوزيعات المتقطّعة الخاصة لأهميتها في الحياة العملية لتفسير كثير من الظواهر التطبيقية ومن التوزيعات التي سوف نتناولها التوزيع المنتظم المتقطّع، محاولة بيرنولي، توزيع ذي الحدين، توزيع فوق الهندسي، وتوزيع بواسون كالتالي.

(٩ ـ ٧ ـ ١) التوزيع المنتظم المتقطّع (uniform distribution)

هو أبسط أنواع التوزيعات المتقطّعة، حيث إن جميع قيم المتغيّر العشوائي لها الاحتمال نفسه وعلى ذلك يكون تعريفه كالتالي:

إذا كان المتغير العشوائي X يأخذ القيم $x_1, x_2, \ldots, x_n$ باحتمال متساوٍ فإن دالة كتلته الاحتمالية كالتالي:

$$f(x) = \begin{cases} \frac{1}{n} & , \quad x = x_1, x_2, \ldots, x_n \\ 0 & , \quad \text{خلاف ذلك} \end{cases} \qquad (٩ - ٦)$$

مثال (٩ ـ ٩)

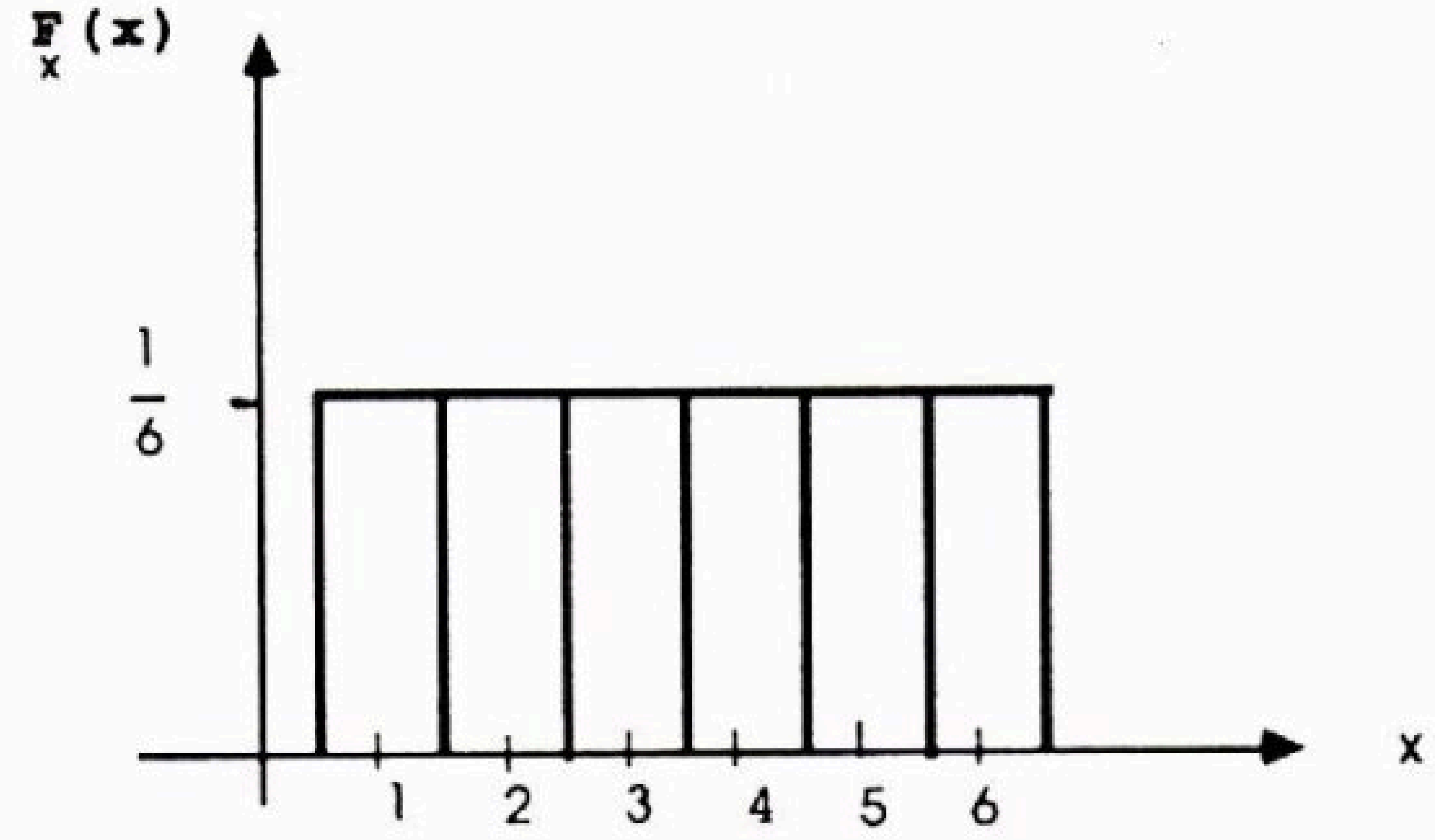
رُمي حجر نرد متجانس مرة واحدة وكان المتغير العشوائي X يمثل الرقم الذي يظهر على الوجه الأعلى. أوجد دالة الكتلة الاحتمالية للمتغير العشوائي X ثم مثلها بيانيًا.

الحـــل

دالة الكتلة $f_X(x)$ هي

$$f_X(x) = \begin{cases} \frac{1}{6} & , \quad x = 1, 2, 3. 4, 5. 6 \\ 0 & , \quad \text{خلاف ذلك} \end{cases}$$

وتمثل بيانيًا في شكل (٩ـ٤) التالي:



شكل (٩ ـ ٤) يمثل المدرج التكراري لدالة الكتلة الاحتمالية للتوزيع المنتظم المتقطع

(٩ ـ ٧ ـ ٢) محاولة برنولي (Bernulli trial)

هناك تجارب كثيرة عشوائية تنتهي إلى نتيجتين اثنتين فقط، تسمى الأولى نجاح (success) ويرمـز لها بالرمز s ، وتسمى الأخرى فشل (failure) ويرمز لها بالرمز f ، وتسمى مثل هذه التجارب بمحاولة بيرنولي. والأمثلة على هذه التجارب كثيرة، نذكر منها على سبيل المثال: عند قذف عملة يكون الناتج هو إما صورة أو كتابة. وعند دراسة صلاحية وحدة من إنتاج مصنع ما تكون إما سليمة أو معيبة. وعند سحب كرة من صندوق يحتوي على كرات بيضاء وحمراء تكون الكرة إما بيضاء أو حمراء. وعند دراسة جنسية عيّنة من العمّال بإحدى المؤسسات بالمملكة يكون العامل إما سعودي أو غير سعودي.

ملاحظـة:

لا يعني الناتج s نجاح أنها بالضرورة هي النتيجة المرغوب فيها.

مما سبق فإن فراغ العيّنة S لمحاولة برنولي هو $S = \{s,f\}$.

وتبعًا لذلك فإننا نعرف متغيرًا عشوائيًا $X(.)$ ذا قيمتين فقط وهما:

$$X(s) = 1 \qquad ; \qquad X(f) = 0$$

وإذا كان $f(1) = p \;;\; f(0) = 1 - p$ فإن دالة الكتلة للمتغيّر X هي :

$$f(x) = \begin{cases} p^x(1-p)^{1-x} \;;\; x = 0,1 \;;\; 0 < p < 1 \\ 0 \quad \text{خلاف ذلك} \end{cases}$$

ويقال إن التوزيع لـ X هو توزيع بيرنولي إذا كانت دالة كتلته الاحتمالية معطاة بالعلاقة السابقة (٩ ـ ٧).

(٩ ـ ٧ ـ ٣) توزيع ذي الحدين (binomial distribution)

نفرض أن تجربتنا العشوائية تتكون من تكرار محاولة بيرنولي عددًا من المرات. وفي المحاولة المكررة فإن احتمال النجاح p (واحتمال الفشل 1 - P) يظل ثابتًا في كل المحاولات.

إذا كانت S هي فضاء العيّنة نجد أن :

$$S = \{(z_1, z_2, \dots, z_n) \;;\; z_i = s \text{ or } f\}$$

سنفرض أيضًا أن هذه المحاولات مستقلة ونحاول إيجاد احتمال بعض الحوادث، إذا كانت $n=5$ مثلاً والمطلوب إيجاد احتمال نتيجة بترتيب معيّن، نأخذ الحادثة A حيث $A=\{(s,s,s,f,f,)\}$ بهذا الترتيب أي الثلاث المحاولات الأولى تعطي نجاحًا والمحاولتين الرابعة والخامسة تعطي فشلاً فيكون :

$$P(A) = p \,.\, p \,.\, p \,.\, (1-p) \,.\, (1-p)$$

$$= p^3 q^2 \quad ; \quad q = 1 - p$$

وبما أن المحاولات مستقلّة، نلاحظ أننا نتوصل إلى النتيجة نفسها لاحتمال حدوث الحادثة التي تتكون من 3 نجاح و 2 فشل وبأي ترتيب آخر. وبالتالي فإن احتمال الحادث $\{(3s, 2f)\}$ بأي ترتيب يساوي :

$$\binom{5}{3} p^3 q^2$$

وعمومًا فإن احتمال الحادثة B حيث $B = \{(xs, (n-x) f)\}$ أو

هو $B = (x \text{ success}, 0 \leqslant x \leqslant n)$

٥

$$P(B) = \binom{n}{x} p^x q^{n-x}$$

فنلاحظ أن احتمال ظهور x نجاح في تكرارنا لمحاولة بيرنولي عدد n مرة يساوي الحد x^{th} في مفكوك ذي الحدين $(p + q)^n$ ومن هذا كان تسميته لهذا التوزيع بالتوزيع ذي الحدين. ويقال إن المتغيّر العشوائي X له توزيع من نوع ذي الحدين إذا كانت دالة كتلته الاحتمالية هي:

$$f_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & ; \ x = 0, 1, 2, \ldots, n \\ 0 & \text{خلاف ذلك} \end{cases} \qquad (٨ - ٩)$$

حيث p هو احتمال النجاح في كل محاولة. وتكتب دالة كتلته الاحتمالية $f_X(x;n,p)$ حيث n ; p عدد المحاولات واحتمال النجاح في المحاولة الواحدة بالترتيب وهما معلمتا هذا التوزيع.

التوقع والتباين لتوزيع ذي الحدين

إذا كان المتغيّر العشوائي X يتبع توزيع ذي الحدين بالمعلمتين n,p فإن دالة كتلته الاحتمالية:

$$f(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x} \ ; \ x = 0, 1, 2, \ldots, n$$

والتوقع للمتغيّر العشوائي X ـ الذي سوف يرمز له بـ $E(X)$ أو μ_X أو μ ـ يعطى كالتالي:

$$\begin{aligned} E(X) &= \sum_x x\, f_X(x;n,p) \\ &= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} \\ &= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} q^{n-x} \\ &= np \sum_{y=0}^{m} \frac{m!}{y!\,(m-y)!} p^y q^{n-x} \end{aligned}$$

$$y = x - 1 \;;\; m = n - 1 \quad \text{حيث}$$
$$= np\,(p + q)^m \quad ; \quad p + q = 1$$
$$= np$$

وكما سبق فإن التباين $V(X)$ للمتغيّر العشوائي X يعطى كالتالي:

$$V(X) = E(X - \mu)^2 = E(X^2) - \mu^2$$

حيث

$$E(X^2) = E[X\,(X - 1) + X] = E[X\,(x - 1)] + E(X)$$

$$E[X\,(X - 1)] = \sum_{x=0}^{n} x\,(x - 1)\,f\,(x, n, p) = n\,(n - 1)\,p^2$$

وذلك باستخدام الخطوات نفسها التي اتبعت في إيجاد $E\,(X)$ فإن

$$E\,(X^2) = n\,(n - 1)\,P^2 + np$$

ومنها ينتج أن تبباين X هو:

$$V\,(X) = \sigma_X^2 = n\,(n - 1)\,P^2 + np - (np)^2$$

$$= np\,(1 - p) = npq$$

أي أن الانحراف المعياري σ_X هو:

$$\sigma_X = \sqrt{npq}$$

فيما يلي بعض الأمثلة لإيجاد التوزيع الاحتمالي والتوقع والتباين لتوزيع ذى الحدين.

مثال (٩ ـ ١٠)

عملة متّزنة رميت ثلاث مرات، فإذا كان المتغيّر العشوائي X يمثل عدد ظهور الصورة أوجد:

ا) دالـة التوزيع الاحتمالي وكذلك التوقع والتباين للمتغير X باستخدام نقاط

فراغ العيّنة.

ب) دالـة التـوزيع الاحتمالي وكذلك التوقع والتباين للمتغيّر X باستخدام دالة الكتلة لتوزيع ذي الحدين.

جـ) قارن بين النتائج المعطاة من (ا)، (ب).

الحـــل

ا) فراغ العيّنة S عند إلقاء قطعة العملة 3 مرات يكون كالتالي:

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

$$n(S) = 2^n = 2^3 = 8$$

وتكون قيم المتغيّر العشوائي X هي

$$x = 0, 1, 2, 3$$

ويمكن حساب الاحتمالات المناظرة لكل قيمة من المتغيّر العشوائي X كالتالي:

$$f_X(0) = P(X = 0) = P(\{TTT\}) = \frac{1}{8}$$

$$f_X(1) = P(X = 1) = P(\{TTH, THT, HTT\}) = \frac{3}{8}$$

$$f_X(2) = P(X = 2) = P(\{HHT, HTH, THH\}) = \frac{3}{8}$$

$$f_X(3) = P(X = 3) = P(\{HHH\}) = \frac{1}{8}$$

ثم نضعها في الجدول التالي:

x	0	1	2	3
f(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

ومن النتائج السابقة يمكن حساب التوقع E(X) والتباين V(X) كالتالي:

$$E(X) = \sum_{x=0}^{3} x\, f(x)$$

$$= (0)\ \frac{1}{8} + (1)\ \frac{3}{8} + (2)\ \frac{3}{8} + (3)\ \frac{1}{8} = \frac{3}{2}$$

وكذلك التباين كما يلي:

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 \\ &= E(X^2) - (\frac{3}{2})^2 \end{aligned}$$

وحيث

$$E(X^2) = \sum_{x=0}^{3} x^2 \,.\, f(x)$$

$$= (0)^2 \,.\, \frac{1}{8} + (1)^2 \,.\, \frac{3}{8} + (2)^2 \,.\, \frac{3}{8} + (3)^2 \,.\, \frac{1}{8} = \frac{24}{8} = 3$$

$$V(X) = 3 - 2.25 = 0.75$$

ويكون الانحراف المعياري σ هو:

$$\sigma = \sqrt{0.75} = 0.866$$

ب) باستخدام دالة الكتلة الاحتمالية لتوزيع ذي الحدين:

$$f_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & , \quad x = 0, 1, 2, 3 \\ 0 & \text{خلاف ذلك} \end{cases}$$

حيث

$$n\ (\text{عدد مرات القذف}) = 3, p = q = \frac{1}{2}$$

$$f_X(0) = \binom{3}{0} (\frac{1}{2})^0 (\frac{1}{2})_3 = \frac{1}{8}$$

$$f_X(1) = \binom{3}{1} (\frac{1}{2})^1 (\frac{1}{2})^2 = \frac{3}{8}$$

$$f_X(2) = \binom{3}{2} (\frac{1}{2})^2 (\frac{1}{2})^1 = \frac{3}{8}$$

$$f_X(3) = \binom{3}{3} (\frac{1}{2})^3 (\frac{1}{2})^0 = \frac{1}{8}$$

وتحسب σ_X^2 و $E(X)$ كما يلي :

$$E(X) = np$$
$$= 3 \times \frac{1}{2} = \frac{3}{2}$$
$$\sigma^2 = npq$$
$$= 3 \times \frac{1}{2} \times \frac{1}{2} = 0.75$$
$$\sigma = \sqrt{npq}$$
$$= \sqrt{0.75} = 0.866$$

وهي النتائج المحسوبة نفسها في الفقرة (ا).

جـ) بالمقارنة بين طريقة الحساب (ا) بواسطة نقاط فراغ العيّنة والحساب (ب) باستخدام دالة الكتلة الاحتمالية للتوزيع ذي الحدين نجد أن الحساب في (ب) أسهل ولا يحتاج لتعيين نقاط فراغ العيّنة وخاصة عندما يكون فضاء العيّنة كبيرًا يصبح معقّدًا. فمثلاً إذا كان عدد مرات قذف قطعة العملة عشر مرات فإن عدد نقاط فضاء العيّنة S يكون $2^n = 2^{10} = 1024$ وإيجاد فضاء العيّنة S يأخذ وقتًا وجهدًا كبيرًا علاوة على احتمالات الخطأ في الحساب بينما التوزيع ذي الحدين لا يتطلب حسابه إيجاد نقاط فضاء العيّنة، وذلك بالتعويض مباشرة في دالة الكتلة الاحتمالية لإيجاد الاحتمال المطلوب وكذلك في الصيغة المبسطة للتوقع والتباين للتوزيع ذي الحدين.

مثال (٩ ـ ١١)

إذا كان 40 % من طلاب إحدى الكليات لا يملكون سيارات فإذا علم أنه في إحدى الشعب 8 طلاب أوجد احتمال أن يكون :

ا) 4 منهم لا يملكون سيارات.

ب) 6 منهم لا يملكون سيارات.

جـ) على الأكثر 2 لا يملكان سيارات .

د) على الأقل 3 لا يملكون سيارات .

الحــل

المتغـيّر العشوائي X يمثل عدد الذين لا يملكون سيارات ويتضح أنه يتوزع توزيعًا ذي حدين .

ويكون قيم المتغير العشوائي X هي :

$$x = 0, 1, 2, 3, \dots, 8$$

ودالة الكتلة الاحتمالية هي :

$$f_X(x) = \begin{cases} P(X = x) = \binom{n}{x} p^x q^{n-x} & ; \quad x = 0, 1, \dots, 8 \\ 0 & \text{خلاف ذلك} \end{cases}$$

حيث

$$n = 8 \quad , \quad p = 0.4 \quad , \quad q = 0.6$$

وعليه فإن الاحتمالات المطلوبة :

أ)

$$f_X(4) = \binom{8}{4} (0.4)^4 (0.6)^{8-4}$$

$$= 0.2322$$

ب)

$$f_X(6) = \binom{8}{6} (0.4)^6 (0.6)^2$$

$$= 0.0413$$

جـ)

$$P(X \leq 2) = f_X(0) + f_X(1) + f_X(2)$$

$$= (0.6)^8 + \binom{8}{1} (0.4)^1 (0.6)^7 + \binom{8}{2} (0.4)^2 (0.6)^6$$

$$= 0.315$$

$$P(X \geqslant 3) = 1 - P(X \leqslant 2)$$

$$= 1 - 0.315$$

$$= 0.685$$

مثال (٩ - ١٢)

وجد في إنتاج أحد المصانع أنه من بين كل 1000 وحدة إنتاج يوجد 150 وحدة معيبة. أخذت عيّنة مكوّنة من 5 وحدات، أوجد الاحتمالات التالية:

ا) الوحدات المختارة كلها سليمة.

ب) على الأكثر توجد وحدة معيبة.

جـ) على الأقل توجد وحدتان معيبتان.

الحــل

إذا كان المتغيّر العشوائي X يمثل عدد الوحدات المعيبة فإنه يتبع توزيع ذي الحدين.

وقيم المتغيّر العشوائي X هي : $x = 0, 1, 2, \ldots, 5$ حيث

$$n = 5 \ , \ p = \frac{150}{1000} = 0.15 \ ; \ q = 0.85$$

$$f(x;n,p) = \binom{n}{x} p^x q^{n-1} \ ; \ x = 0, 1, 2, \ldots , n$$

ا) احتمال الوحدات المختارة كلها سليمة يمكن الحصول عليها بوضع $x=0$ في الدالة السابقة أي أن:

$$f(0) = \binom{5}{0} (0.15)^0 (0.85)^{5-0} = (0.85)^5$$

$$= 0.4437$$

ب) $$P(X \leqslant 1) = f_X(0) + f_X(1)$$

$$= 0.4437 + \binom{5}{1}(0.15)(0.85)^4$$

$$= 0.4437 + 0.3915$$

$$= 0.8352$$

جـ) $$P(X \geq 2) = 1 - p(X \leq 1) = 1 - 0.8325$$

$$= 0.1648$$

(٩ ـ ٧ ـ ٤) التوزيع فوق الهندسي (hypergeometric distribution)

سبق أن تكلمنا عن توزيع ذي الحدين بأنه تكرار لمحاولة برنولي التي من شروطها الاستقلال لكل محاولة عن الأخرى، وكذلك ثبوت احتمال النجاح P لكل المحاولات، وهذا يتحقق في المجتمعات ذات العدد الكبير، مثل نسبة المعيب في الإنتاج، نسبة المدخنين في مجتمع كبير. أو في حالة السحب بإرجاع. فعند السحب بإرجاع يتحقق شرط الاستقلال وثبوت احتمال النجاح P لكل محاولة، وتصبح عملية السحب مثل محاولة برنولي السابقة لتحقق شروطها. أما إذا كان السحب بدون إرجاع وخاصة في المجتمعات الصغيرة، فإن شرط الاستقلال يكون غير محقق، في حالة السحب بدون إرجاع. ولحساب احتمال قيم المتغيّر العشوائي X التي تمثل عدد حالات النجاح، وعلى سبيل المثال: احتمال X يساوي عدد الكرات البيضاء في عيّنة حجمها n ، من صندوق يحتوي على عدد N من الكرات البيضاء والسوداء، وكان سحب العيّنة بدون إرجاع. فإن حساب دالة توزيع الكتلة $f_X(x)$ للمتغيّر العشوائي X لا تتبع توزيع ذي الحدين، لعدم توفر شرط الاستقلال لكل محاولة سحب. وأمثال هذا المتغير يتبع توزيعًا آخر يسمى بالتوزيع فوق الهندسي الذي تعين دالة كتلته الاحتمالية f(x) كما يلي:

نفرض أن لدينا صندوقًا به كرات بيضاء وسوداء عددها N وسحبنا عشوائيًّا منه عيّنة حجمها n كرة. فإذا كان لدينا عدد x كرة بيضاء تمثل حالة النجاح و n – x كرة سوداء تمثل حالة الفشل وكان في الصندوق عدد الكرات البيضاء a وعدد الكرات السوداء b وبذلك يكون:

عدد طرق اختيار x حالة نجاح هو $\binom{a}{x}$ طريقة.
عدد طرق اختيار n − x حالة فشل هو $\binom{b}{n-x}$ طريقة.

حسب قاعدة الضرب عدد طرق اختيار x حالة نجاح و (n − x) حالة فشل هو:

$$\binom{a}{x}\binom{b}{n-x}$$

عدد طرق سحب عينة عشوائية n من مجتمع منتهٍ حجمه a + b هو $\binom{a+b}{n}$ وبذلك يكون احتمال الحصول على x حالة نجاح هو $f_X(x)$ ويكون دالة الكتلة الاحتمالية للتوزيع فوق الهندسي كالتالي:

$$f(x) = P(X=x) = \begin{cases} \dfrac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}} & ; \ \max(0, n-b) \leqslant x \leqslant \min(n,a) \\ 0 & \text{خلاف ذلك} \end{cases} \qquad \text{(٩ - ٩)}$$

ملاحظة

أحيانًا يعطى نسبة إحدى الصفتين p ولتكن الصفة التي عددها a بالنسبة لعدد أفراد المجتمع N . فيمكن حساب a,b كالتالي:

$$a = Np \ ; \ b = Nq \ , \ a + b = N$$

وتكتب دالة الكتلة الاحتمالية للتوزيع فوق الهندسي السابقة على الصورة التالية:

$$f(x) = \begin{cases} \dfrac{\binom{NP}{x}\binom{Nq}{n-b}}{\binom{N}{n}} & ; \ \max(0, n-b) \leqslant x \leqslant \min(n,a) \\ 0 & \text{خلاف ذلك} \end{cases} \qquad \text{(٩ - ١٠)}$$

والتوزيع فوق الهندسي يفيد في المجتمعات التي يمكن تقسيمها إلى صفتين فقط مثل التوزيع ذي الحدين، ولكن هذه المجتمعات تكون محدودة العدد أو صغيرة.

بحيث إن عملية سحب العيّنات بدون إرجاع يؤثر على نسبة إحدى الصفتين، وذلك لصغر حجم المجتمع، والأمثلة على ذلك كثيرة منها، سحب عيّنة من الطلاب في إحدى الشعب، أو عند سحب عيّنة من السيارات من أحد المعارض لدراسة احتمال عدد السيارات المعيبة داخل المعرض، أو سحب عيّنة من السفن بأحد الموانىء لدراسة نسبة السفن الوطنية الموجودة بالميناء، أو سحب عيّنة من الكرات لدراسة احتمال الكرات البيضاء المسحوبة بدون إرجاع من صندوق به كرات بيضاء وغير بيضاء وهكذا.

التوقع والتباين للتوزيع فوق الهندسي

يحسب التوقع $E(x)$ لمتغير عشوائي X يتبع التوزيع فوق الهندسي، مثل ما تم في طريقة حساب التوقع في الحالة العامة ويعوّض عن قيمة $f_X(x)$ بدالة التوزيع فوق الهندسي كما يلي:

$$E(X) = \sum_{x=0}^{n} x\, f_X(x)$$

$$= \sum_{x=0}^{n} x \frac{\binom{NP}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$$

ويكون $E(X) = np$ وذلك بعد إجراء الاختصارات اللازمة.

والتباين $V(X)$ يكون كالتالي:

$$V(X) = E(X^2) - [E(X)]^2$$

$$= \sum_{x=0}^{n} \frac{\binom{NP}{x}\binom{Nq}{n-x}}{\binom{N}{n}} - (np)^2$$

$$= \sum_{x=0}^{n} x(x-1) \frac{\binom{NP}{x}\binom{Nq}{n-x}}{\binom{N}{n}} + \sum_{x=0}^{n} x \frac{\binom{NP}{x}\binom{Nq}{n-x}}{\binom{N}{n}} - (np)^2$$

$$V(X) = npq \left(\frac{N-n}{N-1}\right) = npq \left(1 - \frac{n-1}{N-1}\right)$$

والانحراف المعياري σ هو:

$$\sigma = \sqrt{npq \left(\frac{N-n}{N-1}\right)}$$

فإذا كان حجم المجتمع كبيراً نلاحظ أن:

$$V(X) = npq \quad \text{و} \quad \sigma = \sqrt{npq}$$

وهذا يتحقق مع التباين والانحراف المعياري للتوزيع ذي الحدين أي أن التوزيع فوق الهندسي يؤول إلى التوزيع ذي الحدين عندما يكون حجم المجتمع (N) كبيراً (أنظر تمرين ١١ في نهاية الفصل).

مثال (٩ ـ ١٣)

معرض سيارات به 48 سيارة من بينها 8 سيارات معيبة. اختيرت عينة عشوائية من 5 سيارات أوجد:

ا) دالة الكتلة الاحتمالية والتوقع والتباين لعدد السيارات المعيبة.

ب) احتمال العيّنة كلها سليمة.

جـ) احتمال أن تكون سيارة واحدة معيبة.

د) احتمال أن توجد بها سيارتان معيبتان على الأقل.

الحـــل

نفرض أن المتغير العشوائي X يمثل عدد السيارات المعيبة في العيّنة وهو يتبع التوزيع فوق الهندسي $f_X(x)$ وهو:

$$f_X(x) = \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}} \qquad ; \; x = \max(0, n-b), \ldots, \min(a, n)$$

ا) حيث $N = 48$, $b = 40$, $a = 8$, $n = 5$

فإن

$$f_X(x) = \frac{\binom{8}{x}\binom{40}{5-x}}{\binom{48}{5}} \quad , \; x = 0, 1, 2, 3, 4, 5$$

والتوقع

$$E(x) = np = 5\left(\frac{8}{48}\right) = \frac{5}{6}$$

والتباين

$$W(X) = npq\left(\frac{N-n}{N-1}\right)$$
$$= 5\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)\left(\frac{48-5}{48-1}\right) = 0.64$$

ب) احتمال العيّنة كلها سليمة هي $P(X = 0)$ أي أن:

$$f_X(0) = P(X=0) = \frac{\binom{8}{0}\binom{40}{5-0}}{\binom{48}{5}} = \frac{\binom{40}{5}}{\binom{48}{5}} = 0.38$$

جـ) احتمال أن تكون سيارة معيبة فقط بالعيّنة هو $P(X=1)$ أي أن:

$$f_X(1) = P(X=1) = \frac{\binom{8}{1}\binom{40}{5-1}}{\binom{48}{5}} = \frac{\binom{8}{1}\binom{40}{4}}{\binom{48}{5}} = 0.427$$

د) احتمال أن يوجد على الأقل سيارتان معيبتان هو $P(X\geqslant 2)$ أي أن:

$$P(X \geqslant 2) = 1 - f_X(0) - f_X(1)$$
$$= 1 - 0.38 - 0.427 = 0.193$$

(٩ ـ ٧ ـ ٥) توزيع بواسـون (Poisson distribution)

في بعض المجالات العلمية يكون اهتمامنا الأول إيجاد توزيعات عدد المرات التي تشـاهـد فيهـا ظاهـرة عشـوائية، ولفـترة زمنية محدودة (مسـاحـة أو حجم أو طول

محدود . . الخ) . ومثال ذلك عدد المكالمات الهاتفية التي تصل إلى سنترال مدينة خلال ساعة، عدد الأخطاء الموجودة في إحدى صفحات القاموس، عدد الركاب الذين يصلون إلى محطة حافلات في دقيقة، عدد العضويات (organisms) الموجودة في حجم معين من السائل، عدد حوادث المرور الخطيرة بمدينة خلال أسبوع، أو عدد العيوب في طول قماش في إنتاج ما .

ونحاول الآن إيجاد دالة الكتلة الاحتمالية للمتغير العشوائي X الذي يمثل عدد المرات التي تشاهد فيها هذه الظواهر العشوائية .

ونأخذ مثالاً لذلك عدد المكالمات الهاتفية التي تصل سنترال مؤسسة ما خلال فترة زمنية t . ولنفرض أن المكالمات التي تصل السنترال تحقق الشروط التالية :

ا) احتمال وصول مكالمة واحدة في فترة زمنية قصيرة $(t, t+\triangle t)$ يساوي $\lambda\triangle t+0(\triangle t)$ ، (حيث λ مقدار ثابت يعتمد على طبيعة الظاهرة العشوائية) .

ب) احتمال وصول أكثر من مكالمة في الفترة الزمنية $(t, t+\triangle t)$ صغير جدًا ويمكن اعتباره صفرًا أي أن : $0(\triangle t)=0$.

جـ) عدد المكالمات التي تصل السنترال في الفترات الزمنية المنفصلة تكون مستقلة .

فإذا قسمنا الفترة الزمنية t إلى عدد n فترة زمنية قصيرة $(\triangle t)$ حيث $\triangle t=\frac{t}{n}$ فيمكن اعتبار الفترات الزمنية $\triangle t$ بمثابة عدد من المحاولات المستقلة (محاولات برنولي) وباحتمال ثابت في كل محاولة ويساوي $\lambda\triangle t$.

وباستعمال الشرط (جـ) واستعمال توزيع ذي الحدين نجد أن :

$$P(X=x)=\binom{n}{x}\left(\frac{\lambda t}{n}\right)^{x}\left(1-\frac{\lambda t}{n}\right)^{n-x} \qquad , x=0, 1, \dots , n$$

حيث x يمثل عدد المكالمات التي تصل إلى السنترال في فترة زمنية قدرها t . وبما أن $P(X=x)$ هي قيمة مستقلة عن n فإن :

$$P(X = x) = \lim_{n\to\infty} \binom{n}{x} \left(\frac{\lambda t}{n}\right)^x \left(1 - \frac{\lambda t}{n}\right)^{n-x}$$

$$= e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

ويمكن اعتبار الفترة الزمنية t تساوي وحدة القياس أي أن t=1 فتصبح P(X=x) كالتالي:

$$f_X(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \; ; \; x = 0, 1, 2, \quad \text{.............} \quad (١١ - ٩)$$

وهي ما تسمى بدالة الكتلة الاحتمالية لتوزيع بواسون.

التوقع والتباين لتوزيع بواسون

بالتعويض عن f(x) بدالة الكتلة الاحتمالية لتوزيع بواسون كما يلي:

$$E(X) = \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} = \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda}$$

$$= \lambda e^{\lambda} . e^{-\lambda} = \lambda$$

وبذلك يكون توقع توزيع بواسون يساوي معلمة التوزيع λ .

والتباين V(X) يحسب كالتالي:

$$V(X) = E(X^2) - [E(X)]^2$$

$$\therefore V(X) = \sum_{x=0}^{\infty} x^2 f(x) - (\lambda)^2$$

$$= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} - \lambda^2$$

$$= \sum_{x=0}^{\infty} (x^2 + x - x) \frac{\lambda^x}{x!} e^{-\lambda} - \lambda^2$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x}{x!} e^{-\lambda} + \sum_{i=0} x \frac{\lambda^x}{x!} e^{-\lambda} - \lambda^2$$

$$= \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} e^{-\lambda} + \lambda \sum_{i=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda} - \lambda^2$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

أي أن كل من متوسط وتباين المتغيّر العشوائي الذي يتبع توزيع بواسون يساوي معلمة التوزيع.

مثال (٩ - ١٤)

إذا كان احتمال حدوث ظاهرة ما في فترة صغيرة Δt هو $(0.2)\Delta t$ (وحدة الزمن t هي الدقيقة)، أحسب احتمال رصد هذه الظاهرة عدد مرات قدرها 2 و 1 و 0 مرة في فترة زمنية مقدارها 5 دقائق.

الحــل

نعتبر المتغيّر العشوائي X يساوي عدد مرات رصد الظاهرة وحدة الزمن (دقيقة) وبأخذ القيم .... ,$x = 0, 1, 2$ وله توزيع بواسون

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda} \ , \ x = 0, 1, 2, \ldots\ldots$$

وأما إذا اعتبرنا المتغيّر العشوائي X يساوي عدد مرات رصد هذه الظواهر في الزمن t فيكون X له توزيع بواسون.

$$f(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \ , \ x = 0, 1, 2, \ldots$$

ومن المعطيات نجد أن:

$$\lambda = 0.2 \quad , \quad t = 5 \ , \quad \lambda t = 0.2(5) = 1$$

وعليه فإن

$$f(x) = \frac{(1)^x}{x!} e^{-1} \qquad x = 0, 1, 2, \ldots$$

فيكون

$$f_X(0) = \frac{1}{0!} e^{-1} = e^{-1} = 0.368$$

$$f_X(1) = \frac{1}{1!}\, e^{-1} = e^{-1} = 0.368$$

$$f_X(2) = \frac{1}{2!}\, e^{-1} = 0.184$$

مثال (٩ - ١٥)

إذا كان عدد المرّات التي يتعطل فيها جهاز الكمبيوتر في الأسبوع الواحد يتبع توزيعها التكراري توزيع بواسون حيث $\lambda = 0.4$ فأحسب احتمال أن يعمل الجهاز لمدة أسبوعين بدون توقف.

الحــل

المتغيّر العشوائي X يساوي عدد مرات الأعطال لجهاز الكمبيوتر ويأخذ القيم $X = 0, 1, 2, \ldots$ وله توزيع بواسون

$$f(x) = \frac{\lambda^x}{x!}\, e^{-\lambda} \;,\; x = 0, 1, 2, \ldots$$

احتمال عدم تعطل الجهاز في الأسبوع هو:

$$f(0) = \frac{(0.4)^0}{0!}\, e^{-0.4} = e^{-0.4} = 0.670$$

احتمال عدم تعطل الجهاز لمدة أسبوعين هو:

$$f(0)\,.f(0) = (e^{-0.4})\,(e^{-0.4}) = e^{-0.8} = 0.449$$

مثال (٩ - ١٦)

إذا كان متوسط وصول السفن إلى أحد الموانىء سفينتان في اليوم. أوجد احتمال أن يصل إلى هذا الميناء في يوم معيّن ثلاث سفن.

الحــل

نرمز لعدد السفن بالرمز X وهو متغيّر عشوائي يتبع توزيع بواسون ذات متوسط λ يساوي 2 ويعطى بالعلاقة التالية:

$$f_X(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \ ; \ x = 0, 1, 2, \ldots$$

$$f_X(3) = P(X = 3) = \frac{2^3}{3!} e^{-2} = \frac{8}{6} e^{-2} = 0.18$$

تقريب التوزيع ذي الحدين لتوزيع بواسون

إذا كان المتغيّر العشوائي X يتبع التوزيع ذي الحدين بمتوسط np فإنه عندما تؤول n إلى ما لا نهاية ($n \to \infty$), p تقترب من الصفر ($p \to 0$) فإن توزيع ذي الحدين يؤول إلى توزيع بواسون ذي معلمة λ تعطى بالعلاقة $\lambda = np$ وعليه يكون

$$\lim_{n\to\infty} \binom{n}{x} p^x q^{n-x} = \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \lim_{n\to\infty} \left(1-\frac{1}{n}\right) \ldots \left(1-\frac{x-1}{n}\right) \frac{\lambda^x}{x!} \left(1-\frac{\lambda}{n}\right)^{-x} \left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \left(\left(1+\frac{1}{-n/\lambda}\right)^{-n/\lambda}\right)^{-\lambda}$$

$$= \frac{\lambda^x}{x!} e^{-\lambda}$$

مثال (٩ ـ ١٧)

يحتوي كتاب على 500 صفحة يوجد به 300 خطأ موزعة على صفحات الكتاب. أوجد الاحتمالات التالية:

ا) أن لا تحتوي صفحة معيّنة على خطأ.

ب) أن تحوي صفحة معيّنة خطأ واحدًا فقط.

جـ) أن تحوي صفحة معيّنة خطأين على الأقل.

الحـــل

المتغير العشوائي X يمثل عدد الأخطاء في إحدى الصفحات وهو يتبع التوزيع ذي الحدين بالمعلمتين p و n حيث

$$p = \frac{1}{500} = 0.002 \ , \ n = 300$$

وحيث إن n كبيرة، p صغيرة فإن توزيع X يؤول إلى توزيع بواسون بالمعلمة λ حيث إن

$$\lambda = np$$

$$\therefore \lambda = 300 \text{ x } \frac{1}{500} = 0.6$$

أي أن:

$$f_X(x) = P(X = x) = \frac{\lambda^x}{x\,!} e^{-\lambda}$$

حيث

$$x = 0, 1, 2, \ldots$$

ا) احتمال أن يوجد x = 0 من الأخطاء يعطى

$$f_X(0) = P(X = 0) = \frac{(0.6)^0}{0!} e^{-0.6} = e^{-0.6}$$

$$= 0.55$$

ب) احتمال أن يوجد x = 1 من الأخطاء يعطى

$$f_X(1) = P(X = 1) = \frac{(0.6)}{1!} e^{-0.6} = 0.6e^{-0.6}$$

$$= 0.33$$

جـ) احتمال أن يوجد على الأقل خطأين يعطى

$$P(X \geqslant 2) = 1 - f_X(0) - f_X(1) = 0.12$$

(٩ ـ ٨) نظرية شيبشيف (Chebyshev's theorem)

التباين σ^2 أو الانحراف المعياري σ لأي توزيع احتمالي يقيس مقدار التشتت أو انتشار القيم حول متوسط ذلك التوزيع. فمثلاً إذا كانت قيمة الانحراف المعياري σ

صغيرة فإن احتمال الحصول على قيمة للمتغيّر العشوائي قريبة من متوسط التوزيع μ تكون كبيرة جدًا وهذه الفكرة تعبر عنها **نظريف شيبشف والتي تنص على ما يلي:**

احتمال أن المتغيّر العشوائي X يأخذ قيمة x في المجال المقابل $X(s)$ للمتغير العشوائي X وتنحرف عن المتوسط μ بمقدار $(k\,\sigma)$ على الأكثر فيكون على الأقل هو $(1-\frac{1}{k^2})$.

ويعبر عن ذلك رياضيًا كالتالي:

$$P(|\mu - x| \leqslant k\,\sigma) \geqslant 1 - \frac{1}{k^2}$$

أي أن احتمال الحصول على قيمة x تبعد انحرافين معياريين (2σ) على الأكثر عن المتوسط μ (أي في المجال المقابل داخل حدود $\mu + 2\sigma$ و $\mu - 2\sigma$) يكون على الأقل كالتالي:

$$1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = \frac{3}{4}$$

واحتمال الحصول على قيمة تبعد خمسة انحرافات معيارية على الأكثر عن المتوسط μ (أي في المجال المقابل داخل حدود $\mu + 5\sigma$ و $\mu - 5\sigma$) يكون على الأقل هو $1 - \frac{1}{25}$ ويساوي $\frac{24}{25}$.

واحتمال الحصول على قيمة في المجال المقابل داخل حدود $\mu + 10\sigma$ و $\mu - 10\sigma$) يكون على الأقل هو $(1 - \frac{1}{10^2})$ أي على الأقل $\frac{99}{100}$ وهكذا.

مثال (٩ ـ ١٨)

في 400 رمية لقطعة عملة حصلنا على 128 صورة. هل يمكن القول إن العملة متّزنة؟

الحـــل

إذا كانت العملة متّزنة فإن احتمال الحصول على صورة هو 0.5 ويكون المتغيّر العشوائي X يمثّل عدد الصور ويتبع التوزيع ذي الحدين بمتوسط μ يساوي n p أي أن :

$$\mu = n\,p$$

$$= 400 \times \frac{1}{2} = 200$$

وانحراف معياري σ يساوي $\sigma = \sqrt{n\,p\,q}$ أي أن :

$$= \sqrt{n\,p\,q}$$

$$\sigma = \sqrt{400 \times \frac{1}{2} \times \frac{1}{2}}$$

$$= 10$$

وبـذلك تكون قيمة المتغير العشوائي X التي تمثل 128 صورة تنحرف عن المتوسط μ بمقدار قدره

$$\mu - x = 200 - 128$$

$$= 72$$

ومن نظريف شيبشف نحصل على :

$$72 \leqslant k\,\sigma$$

وحيث $\sigma = 10$ يكون المقدار الثابت k كالتالي :

$$k \geqslant \frac{72}{10}$$

أي أن $k > 7$

أي أن انحراف القيمة 128 عن المتوسط μ بمقدار أكبر من 7 انحرافات معيارية باحتمال أكبر من أو تساوي :

$$1 - \frac{1}{k^2} = 1 - \frac{1}{7^2} = \frac{48}{49} = 0.9796$$

وبذلك فإنه يجب أن نحصل على قيمة في حدود 7σ منحرفة عن المتوسط داخل حدود المجال المقابل في الفترة ما بين (130 ، 270 صورة) وعليه فإنه من العدل أن نقول إن العُملة غير متّزنة. وذلك لأننا حصلنا على 128 صورة وهذا الرقم خارج حدود الفترة السابقة وهي (130, 270).

(٩ ـ ٩) المتغيّر العشوائي المستمر (المتّصل) والتوزيعات المتّصلة

سبق لنا دراسة المتغيّر العشوائي المتقطّع (أو المنفصل) وذكرنا أن المجال المقابل له يكون إما منتهيًا أو غير منتهٍ وقابلاً للعدّ. والمتغيّر العشوائي المتصل (المستمر) هو الذي يكون مجاله المقابل غير منتهٍ وغير قابل للعدّ ويكون عبارة عن جميع القيم داخل فترة (a,b) وتوجد كثير من الظواهر يكون المتغيّر العشوائي الذي يمثلها متصلاً نذكر منها على سبيل المثال ما يلي:

١ ـ طول شخص ما اختير عشوائيًا (لأي جزء من وحدة القياس نختارها).
٢ ـ عمر طالب في شعبة 101 إحص اختير عشوائيًا.
٣ ـ المسافة المقطوعة بسيارة لمقدار جالون من البنزين.
٤ ـ الدخل الشهري لأسرة اختيرت عشوائيًا من أحد البلاد.
٥ ـ درجة حرارة اختيرت عشوائيًا خلال فترة ما بمدينة الرياض.
وهكذا.

وإذا ما تناولنا إحدى الظواهر السابقة كالدخل الشهري للأسر في مدينة ما وكان المتغيّر العشوائي X يمثل هذا الدخل فإننا سوف نجد أسرة ما دخلها الشهري x_1 وأسرة أخرى دخلها الشهري x_2 وسوف توجد أسرة ثالثة دخلها الشهري x_3 يقع بين الدخلين x_1, x_2 مهما كانت القيمتين x_1, x_2 قريبتين من بعضهما وكذلك سوف نجد دخل أسرة رابعة x_4 بين x_1, x_3 وهكذا. وبذلك يكون X المتغيّر العشوائي لدخل الأسرة متغيراً عشوائيًا متصلاً.

ويُعرّف المتغيّر العشوائي المتصل بأنه دالة حقيقية مُعرّفة على الفضاء الاحتمالي (P و ξ و S)*. وفي حالة المتغيّر العشوائي المتصل يجب أن تكون بعض القيود على هذه الـدالـة حتى نضمن كونها مُعرّفة جيدًا وحسنة التصرف. ولتكن X دالة معرفة على (S, ξ, P) عندئذ إذا كانت الحادثة A مجموعة من الأعداد الحقيقية فإن (X in A) تمثل المجموعة المكونة من نقاط العيّنة (w) بحيث X(w) تكون في A . وبشكل خاص نأخذ الحادثة A عبارة عن الفترات التالية:

$$(a \leq X \leq b) = \{w \in S / a \leq X(w) \leq b\}$$

$$(a < X < b) = \{w \in S / a < X(w) < b\}$$

$$(X < b) = \{w \in S / X(w) < b\}$$

وهـكـذا.

والأهم هنا هو أن (X in A) يجب أن يكون حادثة (أي عنصر في حقل سجما ξ عندما تكون الحادثة A فترة، والفترة قد تحوي أولاً نقاطها الطرفية أو قد تمتد من ∞ - إلى ∞) هذه الحقيقة سوف تضمن لنا كون الاحتمالات $P(X \leq b), P(a < X < b), P(a \leq X \leq b)$ . . . الخ، تكون مُعرّفة.

(٩ - ٩ - ١) تعريف المتغيّر العشوائي المتّصل

الدالة الحقيقية X والمُعرّفة على فضاء احتمالي (P و ξ و S) تكون متغيّرًا عشوائيًا متصلاً إذا كانت (X in A) هي حادثة عندما تكون A فترة. وتكون قيم X غير منتهية وغير قابلة للعدّ أي متصلة في هذه الفترة.

(٩ - ٩ - ٢) دالة الكثافة الاحتمالية (probability density function)

لأي متغيّر عشوائي متّصل توجد دالّة تسمى دالة الكثافة الاحتمالية بحيث إن المساحات تحت منحنى هذه الدالّة تعطي الاحتمالات المناظرة للفترات التابعة لها على المحور الأفقي أي أن دالّة الكثافة الاحتمالية إذا تكاملت من a إلى b ($a \leq b$) فإنها تعطي احتمال أن المتغيّر العشوائي سوف يأخذ أي قيمة في الفترة من a إلى b .

* حيث إن S هو فضاء العينة للمتغير العشوائي، ξ حقل سجما معرف عليه القياس الاحتمالي P

تعريف

أي دالة $f_X(x)$ غير سالبة والمُعرَّفة على مجموعة جميع الأعداد الحقيقية تسمى دالة كثافة احتمالية للمتغيّر العشوائي المتصل X إذا وإذا كان فقط:

$$P(a \leq x \leq b) = \int_a^b f_X(x)\, dx$$

لأي أعداد حقيقية ثابتة a و b بحيث $a \leq b$.

لاحظ أنه بما أن X متغيّر عشوائي فإن $(a \leq X \leq b)$ تكون حادثة؛ ذلك لأن (a, b) هي فترة.

ملاحظة

$f_x(c)$ قيمة دالة الكثافة الاحتمالية للمتغيّر العشوائي X عند النقطة (C) لا تعطي الاحتمال $P(X=c)$ كما في حالة المتغيّر العشوائي المنفصل؛ لأن في حالة المتغيّر العشوائي المتصل الاحتمالات تعطى دائمًا بتكاملات للدالة $f_X(x)$ في فترات وفي حالتنا تكون الفترة $(c \leq x \leq c)$ حادثة خالية وبذلك يكون $P(X=c) = 0$ لأي ثابت حقيقي (c) .

ونلاحظ أيضًا أن دالة الكثافة الاحتمالية للمتغيّر العشوائي X لا تتأثر بإضافة أو عدم إضافة نقط النهايات للفترة في a إلى b أي أن:

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$$

نظرية (٩ ـ ١)

أي دالة يمكن أن تكون دالة كثافة احتمالية لمتغيّر عشوائي متّصل X إن كانت قيمتها $f(x)$ تحقق الشرطين التاليين:

$$1)\ f_X(x) \geq 0 \qquad \text{for} \quad -\infty < x < \infty$$

$$2)\ \int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

مثال (٩ - ١٩)

دالة الكثافة الاحتمالية للمتغيّر العشوائي المتّصل X تعطى بالعلاقة :

$$f_X(x) = \begin{cases} k\,e^{-3x} & \text{for} \quad x > 0 \\ 0 & \text{خلاف ذلك} \end{cases}$$

أوجد قيمة k وكذلك ما هو احتمال X تأخذ قيم في الفترة (1 , 0.5) .

الحـــل

لإيجاد قيمة k نحل المعادلة التالية مستخدمين الشرط الثاني في النظرية السابقة .

$$\int_{-\infty}^{\infty} f_x(x)\,dx = \int_{0}^{\infty} k\,e^{-3x}\,dx = k \lim_{x\to\infty} \frac{e^{-3x}}{-3} \Big|_{0}^{\infty}$$

$$= \frac{k}{3} = 1$$

فيكون قيمة k = 3 .

ولحساب الاحتمال في الفترة (0.5, 1) كالتالي :

$$P(0.5 < X < 1) = \int_{0.5}^{1} 3e^{-3x}dx = -e^{-3x} \Big|_{0.5}^{1}$$

$$= -e^{-3} + e^{-1.5} = 0.173$$

(٩ - ٩ - ٣) دالة التوزيع التراكمي (cumulative distribution function)

إذا كان X متغيّرًا عشوائيًا متصلاً فإن الدالة المعطاة بالعلاقة التالية :

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(t)dt \;\; ; \; -\infty < x < \infty$$

حيث $f_X(t)$ هي قيمة دالة الكثافة الاحتمالية للمتغيّر العشوائي X عند t فإن $F_X(x)$ تسمى دالة التوزيع أو التوزيع التراكمي للمتغير العشوائي X ، وخواص دالة التوزيع التراكمي قد سبق ذكرها في حالة المتغيّر العشوائي المنفصل .

نظرية (٩ ـ ٢)

إن كان $f_X(x)$ ، $F_X(x)$ هما على الترتيب قيمتا دالة الكثافة الاحتمالية ودالة التوزيع التراكمي للمتغيّر العشوائي X عند x فإن

$$P(a \leqslant X \leqslant b) = F_X(b) - F_X(a)$$

لأي ثوابت حقيقية a و b تحقق $a \leqslant b$

$$f_X(x) = \frac{dF_{X}(x)}{dx}$$ أيضًا وذلك إن وجدت مشتقة $F_X(x)$.

كذلك تكون

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

مثال (٩ ـ ٢٠)

أوجد دالة التوزيع التراكمي للمتغيّر العشوائي المتّصل X الذي أعطيت دالة كثافته الاحتمالية في مثال (٩ ـ ١٩) السابق ومستعملاً دالة التوزيع التراكمي لإيجاد $P(0.5 \leqslant X \leqslant 1)$ وقارنها بالحل في مثال (٩ ـ ١٩) السابق.

الحـــل

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{0}^{x} 3e^{-3t}\, dt = -e^{-3t}\Big|_{0}^{x}$$
$$= 1 - e^{-3x}$$

وتكتب على الشكل

$$F_X(x) = \begin{cases} 0 & \text{for } \quad x \leqslant 0 \\ 1 - e^{-3x} & \text{for } \quad x > 0 \end{cases}$$

لإيجاد $P(0.5 \leqslant X \leqslant 1)$ كالتالي :

$$P(0.5 \leqslant X \leqslant 1) = F_X(1) - F_X(0.5) = -e^{-3} + e^{-1.5} = 0.173$$

وهي النتيجة نفسها التي سبق الحصول عليها في مثال (٩ ـ ١٩) السابق.

مثال (٩ ـ ٢١)

أوجد دالة الكثافة الاحتمالية للمتغيّر العشوائي الذي له دالة التوزيع التراكمي المعطاة بالعلاقة:

$$F_X(x) = \begin{cases} 0 & \text{for} \quad x \leq 0 \\ x & \text{for} \quad 0 < x < 1 \\ 1 & \text{for} \quad x \geq 1 \end{cases}$$

الحـــل

$$f_X(x) = \frac{d\,F_X(x)}{dx} = \begin{cases} 1 & \text{for} \quad 0 < x < 1 \\ 0 & \text{خلاف ذلك} \end{cases}$$

(٩ ـ ٩ ـ ٤) التوقع والتباين للمتغيّر العشوائي المتّصل

يمكن حســاب التـوقع والتباين للمتغيّر العشوائي المتّصل X كما سبق حسابه للمتغيّر العشوائي المتقطّع وذلك باستبدال العلامة Σ التي ترمز للمجموع بالعلامة ∫ التي ترمز للتكامل وبذلك يكون التوقع E(x) الذي يرمز له بالرمز μ كالتالي:

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

والتباين V(X) أو (σ^2) كالتالي:

$$V(X) = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$$

ومن خواص التوقع يمكن كتابة V(X) بشكل مبسط كما يلي:

$$V(X) = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx - \mu^2$$

وسوف نقوم بتوضيح طريقة حساب كل من التوقع والتباين للمتغيّر العشوائي المتّصل من المثال التالي.

مثال (٩ ـ ٢٢)

إذا كانت الدالة التالية

داخل الفترة $0 \leqslant X \leqslant 3$ تكون

$$f_X(x) = \frac{2}{9}\ x$$

وخلاف ذلك تكون

$$f_X(x) = 0$$

ا) اثبت أنها دالة كثافة احتمالية للمتغيّر العشوائي X .

ب) أحسب قيمة احتمال أن يقع المتغيّر العشوائي داخل الفترة $(2, \frac{5}{2})$.

جـ) أوجد التوقع μ والانحراف المعياري σ .

الحــل

ا) لكي تكون الدالة $f_X(x)$ دالة كثافة يجب أن تحقق الشرطين:

$$\text{(i)}\ f(x) \geqslant 0 \quad , \quad \text{(ii)}\ \int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

من تعريف الدالة نجد أن $f_X(x)$ دائمًا موجبة أي أن الشرط (1) قد تحقق.
وكذلك

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_0^3 \frac{2}{9}\ x\, dx = \frac{1}{9}\ (x^x)\Big|_0^3$$

$$= \frac{1}{9}\ (9-0)\ =\ 1$$

أي أن الشرط الثاني قد تحقق وبذلك فهي دالة كثافة احتمالية.

ب) احتمال أن يقع المتغيّر العشوائي في الفترة $(2, \frac{5}{2})$ يعطي

$$P(2 \leq X \leq \frac{5}{2}) = \int_{2}^{5/2} \frac{2}{9}\, x\, dx$$

$$= \frac{1}{9}\,(x^2)\Big|_{2}^{5/2} = \frac{1}{9}\,(\frac{25}{4} - 4) = \frac{1}{4}$$

جـ) يحسب المتوسط كما يلي:

$$\because\ \mu = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

$$\therefore\ \mu = \int_{0}^{3} \frac{2}{9}\, x^2\, dx = \frac{2}{27}\,.\, x^3\ \Big|_{0}^{3} = 2$$

ويحسب التباين σ^2 كما يلي:

$$\because\ \sigma^2 = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx - \mu^2$$

$$\therefore\ \sigma^2 = \int_{0}^{3} \frac{2}{9}\, x^3\, dx - (2)^2$$

$$= \frac{1}{18}\, x^4\ \Big|_{0}^{3} - 4$$

$$= \frac{81}{18} - \frac{72}{18} = \frac{9}{18} = \frac{1}{2}$$

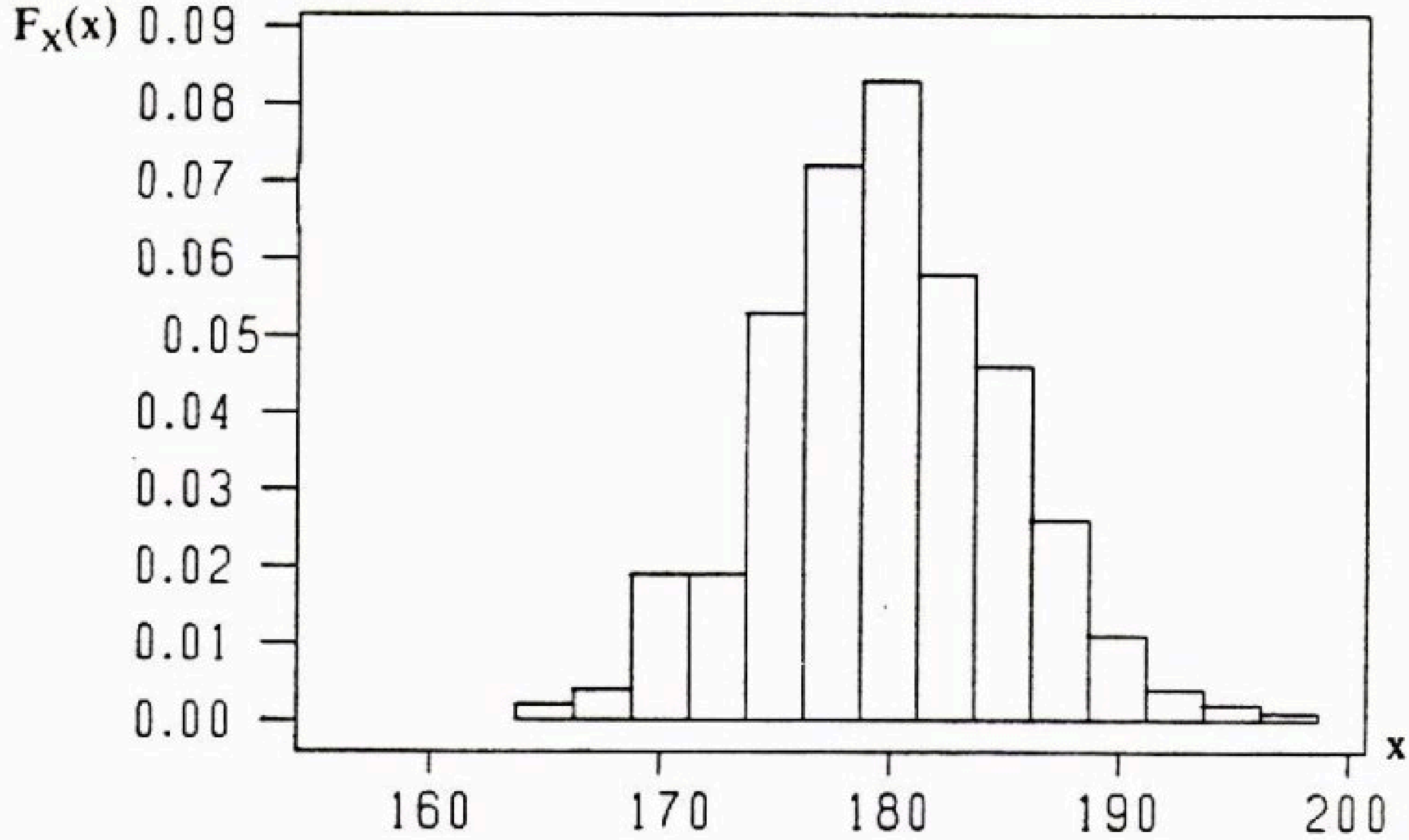
الانحراف المعياري σ يكون

$$\sigma = \sqrt{0.5} = 0.7$$

(٩ ـ ١٠) التوزيع الطبيعي

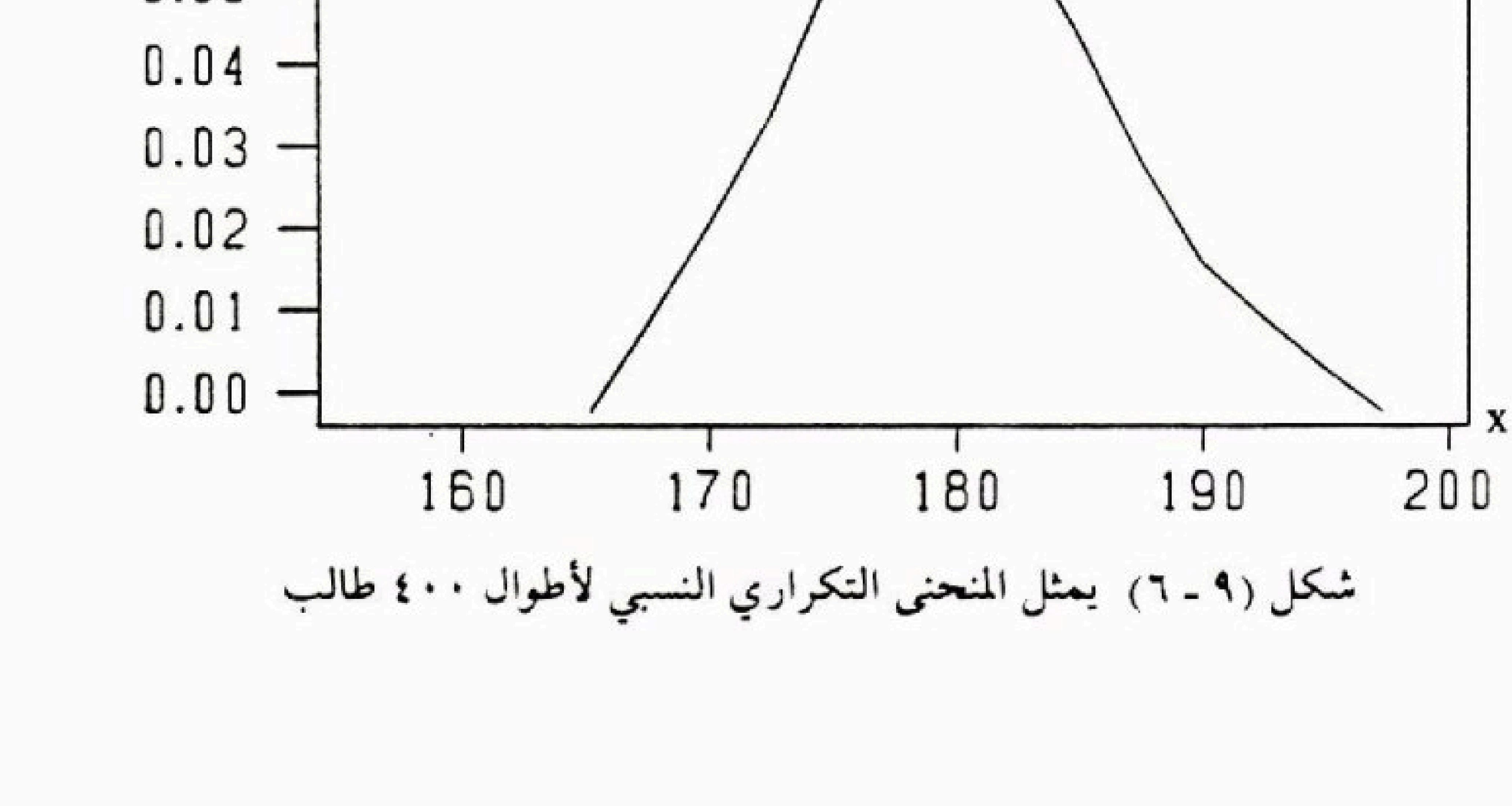
Natural distribution

إذا أخذنا على سبيل المثال أطوال مجموعة من 400 طالب من طلاب جامعة الملك سعود ولخصت هذه الأطوال في توزيع تكراري نسبي فإنه يمكن رسم المدرج التكراري النسبي شكل (٩ ـ ٥) التالي:



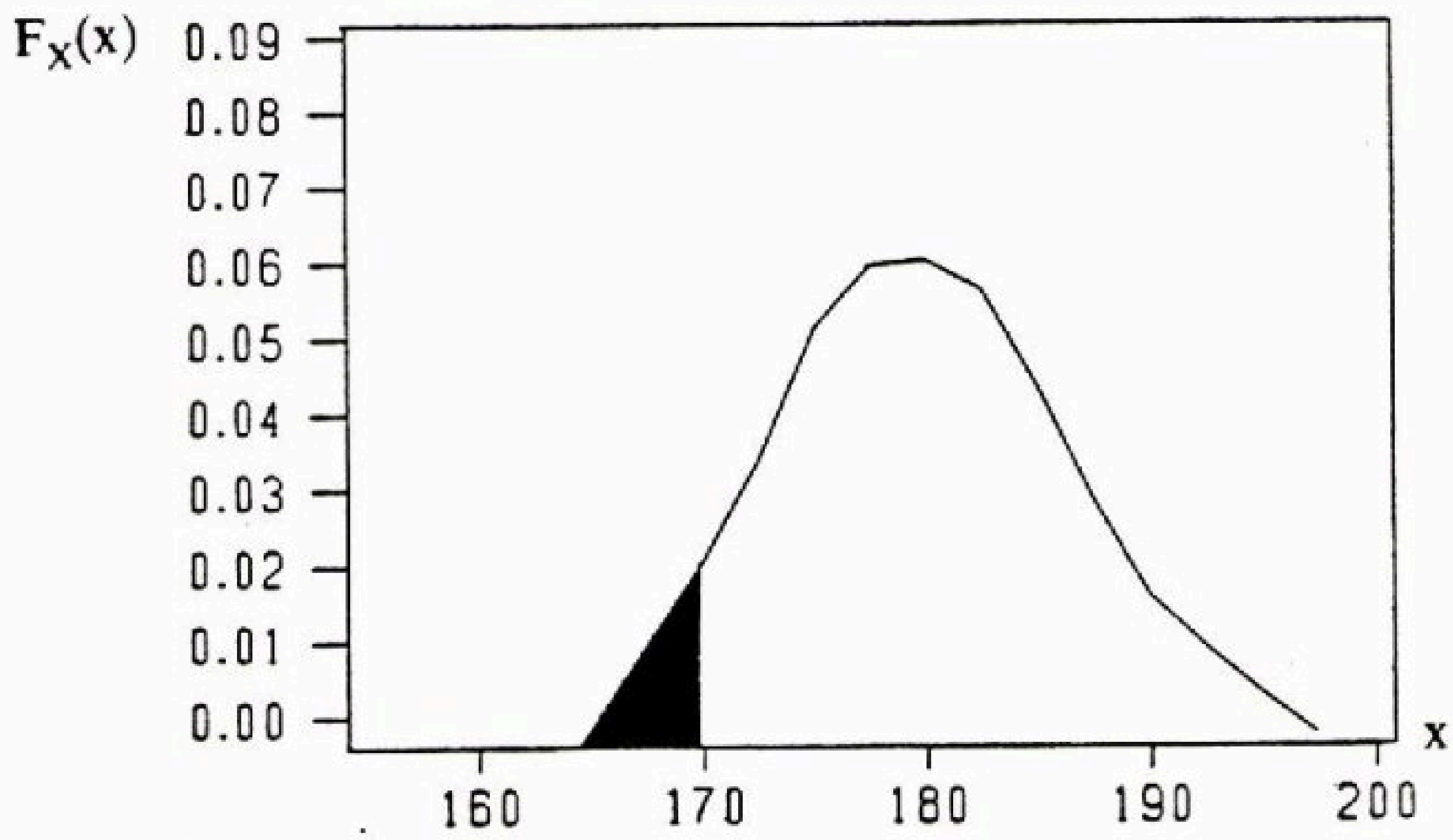
شكل (٩ ـ ٥) يمثل المدرج التكراري النسبي لأطوال ٤٠٠ طالب

ومن الشكل نرى أن جميع أطوال الطلبة تتراوح بين 155 سم، 196 سم ولكن ليس من السهل التعامل مع التوزيع التكراري النسبي السابق ولهذا فإن الإحصائيين يقومون بتقريب التوزيع السابق بمنحنى أملس، شكل (٩ ـ ٦) التالي:



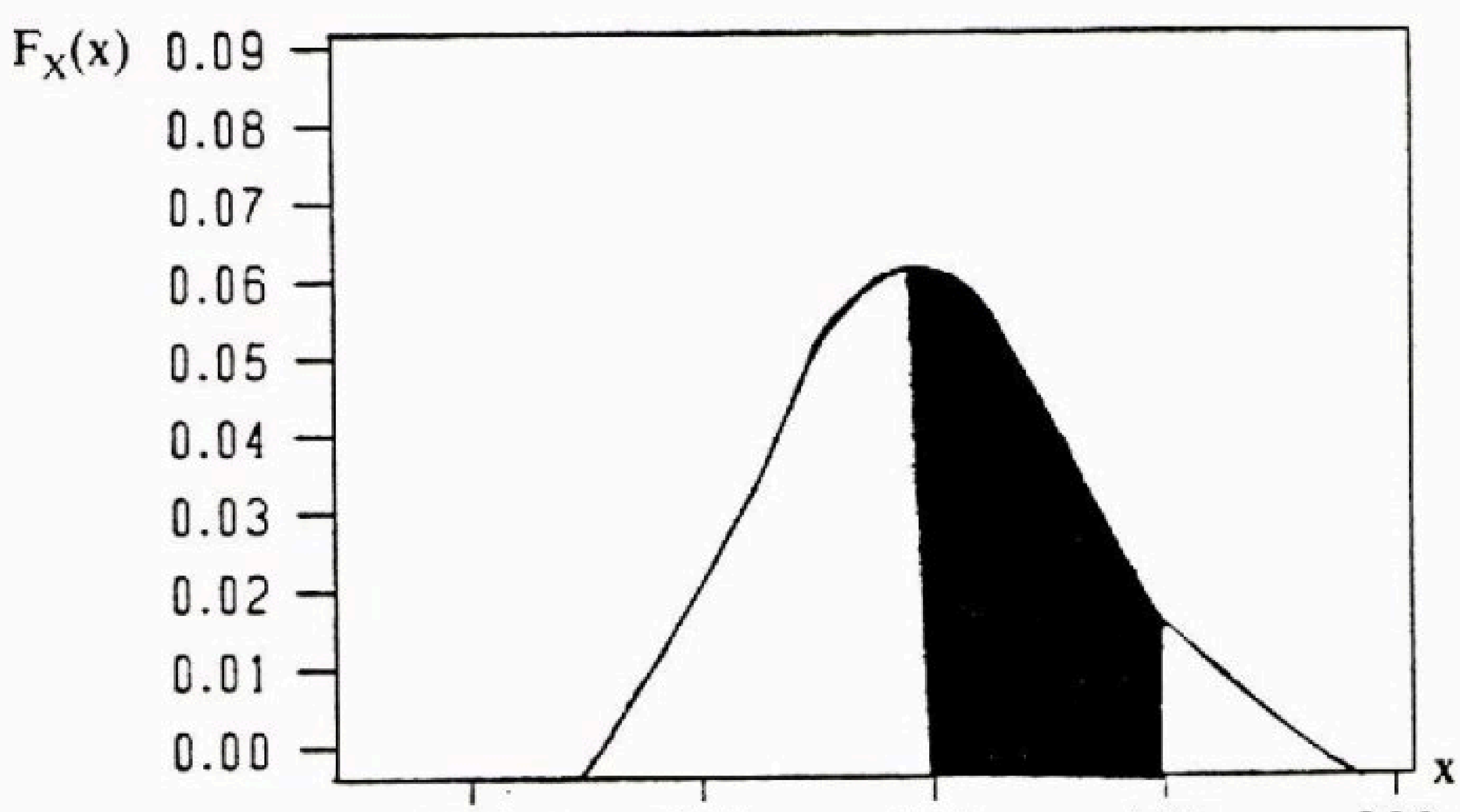
شكل (٩ ـ ٦) يمثل المنحنى التكراري النسبي لأطوال ٤٠٠ طالب

من مميزات منحنى التوزيع التكراري النسبي وأمثاله أن المساحة تحت المنحنى تمثل احتمالات كما سبق أن تكلمنا في الحالة العامة لتوزيع المتغيّر العشوائي المتّصل. فمثلاً إذا سحبنا قيمة للمتغير العشوائي X تساوي 160 فإن $P(X \leq 160)$ يعطى بالمساحة المظللة في شكل (٩ ـ ٧) التالي:



شكل (٩ ـ ٧) يمثل نسبة الطلّاب الذين أطوالهم أقل أو يساوي ١٦٠سم

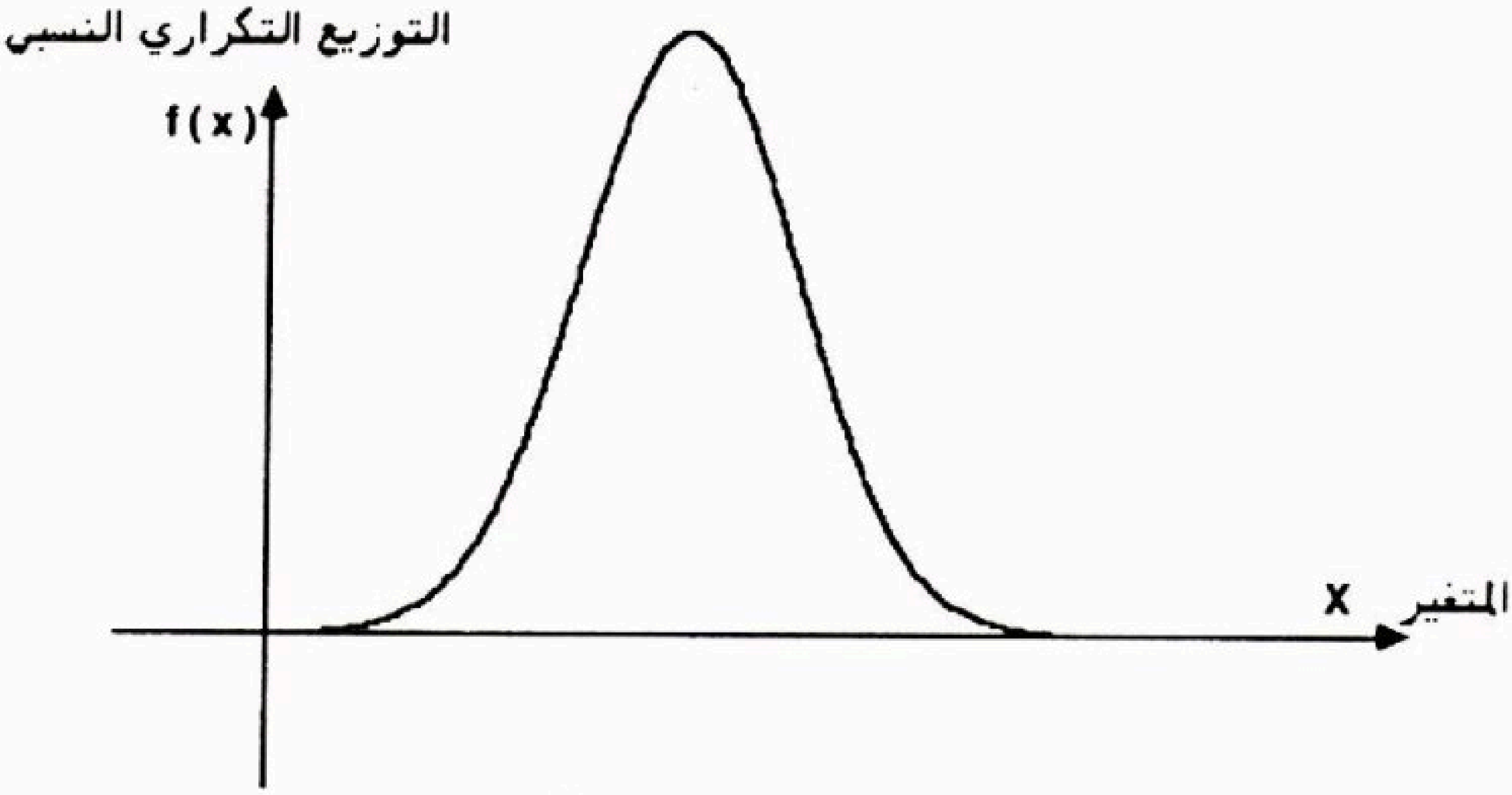
وبالمثل يمكن إيجاد احتمال المتغير العشوائي X داخل فترة (180, 190) من المنحنى التكراري النسبي السابق بالشكل المظلّل في شكل (٩ ـ ٨) التالي:



شكل (٩ ـ ٨) يمثل نسبة الطلّاب الذين أطوالهم بين ١٨٠سم/ ١٩٠سم

ويجب أن ننوه بأنه في حالة المتغيرات العشوائية المتصلة (المستمرة) فإنه بإمكاننا الحصول على احتمالات لقيم المتغير العشوائي الواقع داخل فترة فقط حيث إن احتمال المتغيّر العشوائي المتّصل عند قيمة معيّنة يساوي صفرًا أي أن $P(X=158)=0$ مثلاً؛ وذلك لعدم إمكان تحديد طول شخص يساوي 158 سم تمامًا. وبذلك يكون من المفيد جدًا معاملة بعض المتغّيرات العشوائية المتقطّعة (المنفصلة) على أنها متغيّرات عشوائية مستمرة وذلك في حالة أخذ المتغيّر العشوائي المتقطّع لعدد كبير من القيم الممكنة (مثلاً رمي قطعة عملة من 100 إلى 1000 مرة).

والمنحنى في شكل (٩ ـ ٩) التالي يمثل متغيرًا عشوائيًا X يمثل أوزان مجتمع ما.



شكل (٩ ـ ٩) يمثل منحنى التوزيع الطبيعي

وفي علم الإحصاء تم تطوير عدد من المنحنيات مثل المنحنى المتماثل السابق وهو يمثل تقريبًا جيدًا للعديد من المدرجات التكرارية النسبية للعديد من المتغيّرات العشوائية المتّصلة، ومن أهم هذه التوزيعات التوزيع الطبيعي.

والتوزيع الطبيعي هو عبارة عن توزيع له شكل الجرس المقلوب وتأتي أهمية هذا التوزيع لكونه يقارب بشكل جيد توزيعات الكثير من القياسات المختلفة التي منها

الأطوال، الأوزان، درجات الامتحان، ضغط الدم، حجم الرأس، عمر إطار سيارة، أخطاء القياسات. . . الخ. وتعطى دالة الكثافة $f_X(x)$ للمتغيّر العشوائي X الذي يتبع التوزيع الطبيعي ويكون مجاله المقابل X(s) هو مجموعة الأعداد الحقيقية R حيث $R = (-\infty , \infty)$ كالتالي:

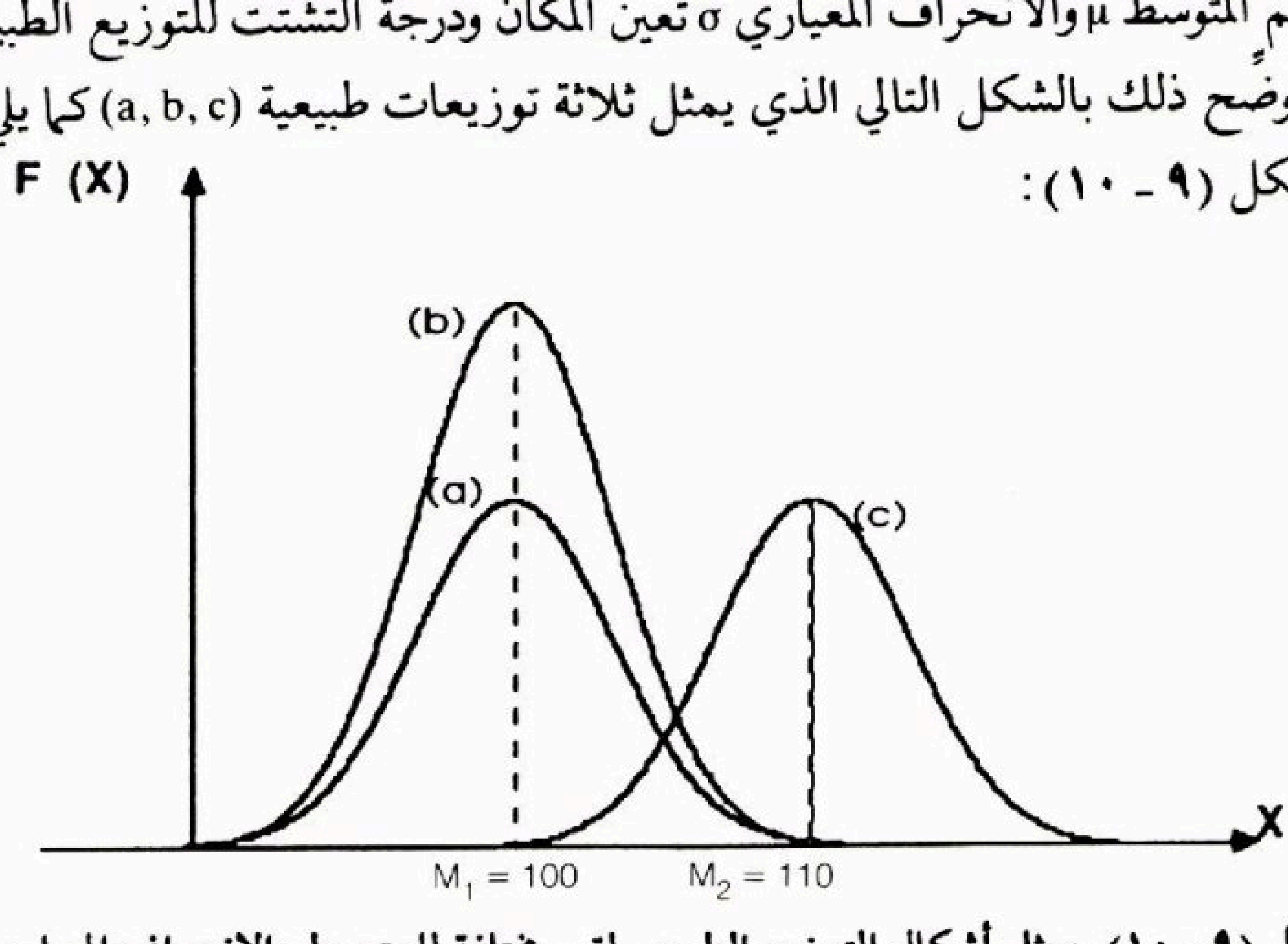
$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \text{(٩ - ١٢)}$$

حيث

$$\sigma > 0 \ , \ -\infty < \mu < \infty \ , \ -\infty < X < \infty$$

(٩ - ١٠ - ١) بعض خواص التوزيع الطبيعي

١) قيم المتوسط μ والانحراف المعياري σ تُعيِّن المكان ودرجة التشتت للتوزيع الطبيعي وتُوضِّح ذلك بالشكل التالي الذي يمثل ثلاثة توزيعات طبيعية (a, b, c) كما يلي في شكل (٩ - ١٠):



شكل (٩ - ١٠) يمثل أشكال التوزيع الطبيعي لقيم مختلفة للمتوسط والانحراف المعياري

التوزيعان (a), (b) لهما المتوسط نفسه $\mu_1 = 100$ ولكن الانحرافين المعياريين مختلفان. والتوزيعان (a), (c) لهما الانحراف المعياري نفسه ولكن يختلفان في

المتوسطين حيث $\mu_1 = 100$, $\mu_2 = 110$ والتوزيعين (c) و (b) غير متساويين في كل من المتوسط والانحراف المعياري .

٢) التوزيع الطبيعي متناظر حول المتوسط μ .

٣) منحنى التوزيع الطبيعي يقترب طرفاه أكثر فأكثر من المحور الأفقي ولكن لا يمسه أو يقطعه أبدًا .

٤) المساحة الكلية تحت المنحنى تساوي واحدًا صحيحًا .

٥) % 68 من مساحة المنحنى الطبيعي تقريبًا تنحصر في الفترة من $\mu - \sigma$ إلى $\mu + \sigma$.

٦) % 95 من المساحة تحت المنحنى تنحصر داخل الفترة من $\mu - 2\sigma$ إلى $\mu + 2\sigma$.

ونرمز للمتغيّر العشوائي الموزع توزيعًا طبيعيًّا بالرمز

$$X \sim N(\mu, \sigma^2)$$

حيث μ هي متوسط المتغيّر العشوائي X و σ هي الانحراف المعياري للمتغير العشوائي X . وهكذا نقول إن X له توزيع طبيعي بمتوسط μ وانحراف معياري σ .

مثال (٩ ـ ٢٣)

إذا علم أن مقياس الذكاء X في مجتمع معين موزع توزيعًا طبيعيًّا ذا متوسط 100 وانحراف معياري 10 .

مثّل الاحتمالات التالية بمساحات مظلّلة تحت منحنى التوزيع الطبيعي :

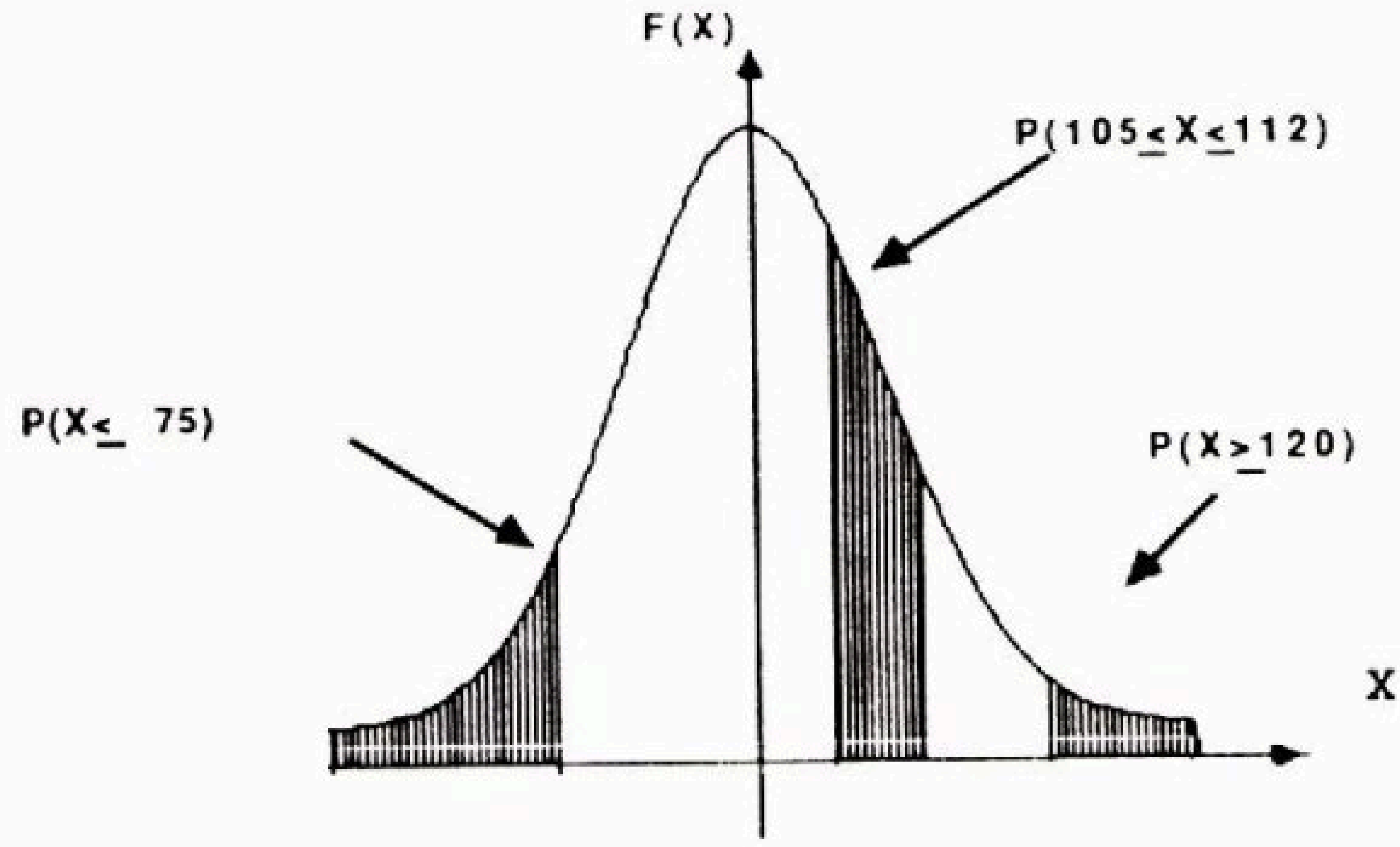
$$P(X \leqslant 75)\ ,\ P(105 < X \leqslant 112)\ ,\ P(X \geqslant 120)$$

الحـــل

نرسم المنحنى للتوزيع الطبيعي N(100,100) كما يلي في شكل (٩ ـ ١١) .

ويمكن أن نرمز لقيمة الاحتمال $P(X \leqslant x)$ بالدالة $F_X(x)$ وتسمى بدالة الاحتمال التراكمية وتعطى للتوزيع الاحتمالي بالعلاقة التالية :

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(t)\, dt$$



شكل (٩ - ١١) يمثل قيم الاحتمالات المطلوبة في المثال (٩ - ٢٣)

وبالتعويض عن دالة كثافة الاحتمال $f_X(x)$ للتوزيع الطبيعي نحصل على

$$f_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

وبالتالي تكون المساحات المظللة في الشكل السابق هي التكاملات التالية:

$$F(75) = \int_{-\infty}^{75} \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-100}{10}\right)^2} dt$$

$$P(105 \leqslant X \leqslant 112) = \int_{105}^{112} \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-100}{10}\right)^2} dt$$

$$P(X \geqslant 120) = \int_{120}^{\infty} \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-100}{10}\right)^2} dt$$

والتكـامـلات السابقة عبارة عن الاحتمالات المطلوبة وهي المساحات المظلّلة بالشكل السابق، شكل (٩ - ١١).

(٩ - ١٠ - ٢) التوزيع الطبيعي القياسي (standard normal distribution)

كما بيّنا أن التوزيع الطبيعي للمتغيّر العشوائي X يعتمد على متوسطه μ وانحرافه المعياري σ أي $N(\mu,\sigma^2)$ وهاتان المعلمتان تؤثران على مكان التوزيع وعلى مقدار تشتته فكل قيمة للمتوسط μ وكل قيمة للتباين σ يتغير شكل ومكان التوزيع. ولكن يوجد توزيع طبيعي واحـد ثابت ومجدول يمكن تحويل جميع التوزيعات الطبيعية الأخرى إليه. ويسمى التوزيع الطبيعي القياسي. وهو توزيع طبيعي ذو متوسط صفر وانحراف معياري يساوي واحدًا صحيحًا وعادة ما يرمز للمتغيّر العشوائي الذي يتبع التوزيع الطبيعي القياسي بالرمز Z أي أن $Z \sim N(0,1)$ ودالة الكثافة الاحتمالية للمتغيّر العشوائي Z يرمز لها بالرمز $\phi(z)$ وتعطى بالعلاقة

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \ldots\ldots\ldots \quad (٩ - ١٣)$$

حيث $\infty > z > \infty -$ أي المجال المقابل لها $Z(S)$ هو مجموعة الأعداد الحقيقية R في الفـترة $(-\infty , \infty)$ وتكـون دالـة احتماله التراكمية $\Phi(z)$ هي المساحة تحت المنحنى الطبيعي القياسي من $\infty -$ إلى النقطة z ويعبر عن ذلك رياضيًا كما يلي:

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt \quad \ldots\ldots \quad (٩ - ١٤)$$

وبكلّما توجد في الحياة العملية أية ظاهرة طبيعية يكون متوسطها صفرًا وانحرافها المعياري واحدًا ولكن جميع القيم غير القياسية يمكن تحويلها إلى قيم قياسية باستخدام العلاقة:

$$Z = \frac{X - \mu}{\sigma}$$

فمثلاً إذا كان لدينا المتغيّر العشوائي $X \sim N(16, 16)$ يكون المتغيّر $\frac{X-16}{4}$ أيضًا متغيرًا عشوائيًا طبيعيًا ولكن هذا المتغير له متوسط $\mu=0$ وانحراف معياري $\sigma=1$ أي أن

$$Z = \frac{X-16}{4}$$

ويكون المتغيّر العشوائي Z موزعًا توزيعًا طبيعيًا قياسيًا ويسمى بالمتغيّر العشوائي القياسي. ومن أهم خصائص المنحنى الطبيعي القياسي أنه منحنى متناظر حول المحور الرأسي والمساحة الكلية المحصورة بينه وبين المحور الأفقي تساوي واحدًا وهي موزعة كالتالي:

% 68 من المساحة تحت المنحنى تنحصر بين (1 , 1-) ،
% 95 من المساحة تحت المنحنى تنحصر بين (2 , 2-) ،
% 99.75 من المساحة تحت المنحنى تنحصر بين (3 , 3-) .

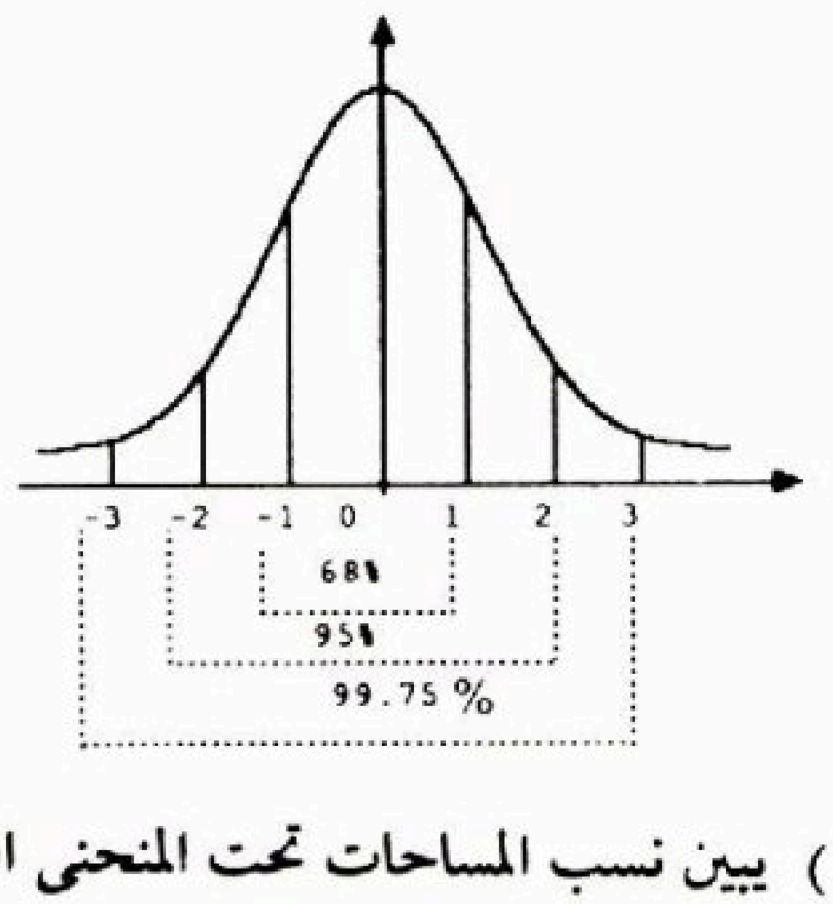
وهذه المساحات السابقة تكتب بصيغة الاحتمالات كالتالي:

$$P(-1 \leqslant Z \leqslant 1) = 0.68$$

$$P(-2 \leqslant Z \leqslant 2) = 0.95$$

$$P(-3 \leqslant Z \leqslant 3) = 0.9975$$

وتمثل الاحتمالات السابقة باستخدام المنحنى الطبيعي القياسي في شكل (٩ - ١٢) التالي:



شكل (٩ - ١٢) يبين نسب المساحات تحت المنحنى الطبيعي القياسي

مثال (٩ ـ ٢٤)

إذا علم أن أطوال مجموعة من الطلّاب يتبع التوزيع الطبيعي بمتوسط مقداره 165 سم وانحراف معياري 5 سم، أوجد القيم المعيارية لأطوال الطلّاب التالية:

150, 158, 167, 172

الحـــل

القيم المعيارية z تعطى بالعلاقة التالية:

$$z = \frac{x - \mu}{\sigma}$$

حيث $\mu = 165$ ، $\sigma = 5$

z_1 القيمة المعيارية للطالب الأول حيث $x_1 = 150$ تكون كالتالي:

$$z_1 = \frac{150 - 165}{5} = -3$$

z_2 القيمة المعيارية للطالب الثاني حيث $x_2 = 158$ تكون كالتالي:

$$z_2 = \frac{158 - 165}{5} = -1.4$$

z_3 القيمة المعيارية للطالب الثالث $x_3 = 167$ تكون كالتالي:

$$z_3 = \frac{167 - 165}{5} = 0.4$$

z_4 القيمة المعيارية للطالب الرابع $x_4 = 172$ تكون كالتالي:

$$z_4 = \frac{172 - 165}{5} = 1.4$$

مثال (٩ ـ ٢٥)

باستخدام البيانات التي في مثال (٩ ـ ٢٤) أوجد الأطوال الحقيقية للطلّاب إذا علم أن أطوالهم القياسية هي:

$-0.52, 0.82, -0.8, 1.2$

الحـــل

حيث إن القيم المعيارية z تعطى بالعلاقة التالية :

$$z = \frac{x - \mu}{\sigma}$$

فإن القيمة الحقيقية x تعطى من العلاقة السابقة كالتالي :

$$x = \mu + \sigma z$$

فيكون القيمة الحقيقية x_1 للطالب الأول حيث $z_1 = 1.2$ تكون كالتالي :

$$x_1 = 165 + 5\,(1.2) = 171 \quad \text{سم}$$

x_2 القيمة الحقيقية للطالب الثاني حيث $z_2 = -0.8$ تكون كالتالي :

$$x_2 = 165 + 5\,(-0.8) = 161 \quad \text{سم}$$

x_3 القيمة الحقيقية للطالب الثالث حيث $z_3 = 0.82$ تكون كالتالي :

$$x_3 = 165 + 5\,(0.82) = 169.1 \quad \text{سم}$$

x_4 القيمة الحقيقية للطالب الرابع حيث $z_4 = -0.52$ تكون كالتالي :

$$x_4 = 165 + 5\,(-0.52) = 162.4 \quad \text{سم}$$

مثال (٩ - ٢٦)

المتغير العشوائي Z موزع توزيعًا طبيعيًا قياسيًا. باستخدام الجدول الإحصائي الذي يعطي المساحة تحت المنحنى الطبيعي القياسي. أوجد الاحتمالات

$$\text{(i)}\ P(Z \leq 1.72)\ ;\ \text{(ii)}\ P(Z \leq -0.54)\ ;\ \text{(iii)}\ P(Z \leq 1.07)$$

$$\text{(iv)}\ P(Z \geq 0.29)\ ;\ \text{(v)}\ P(-1.91 \leq Z \leq 0.45)$$

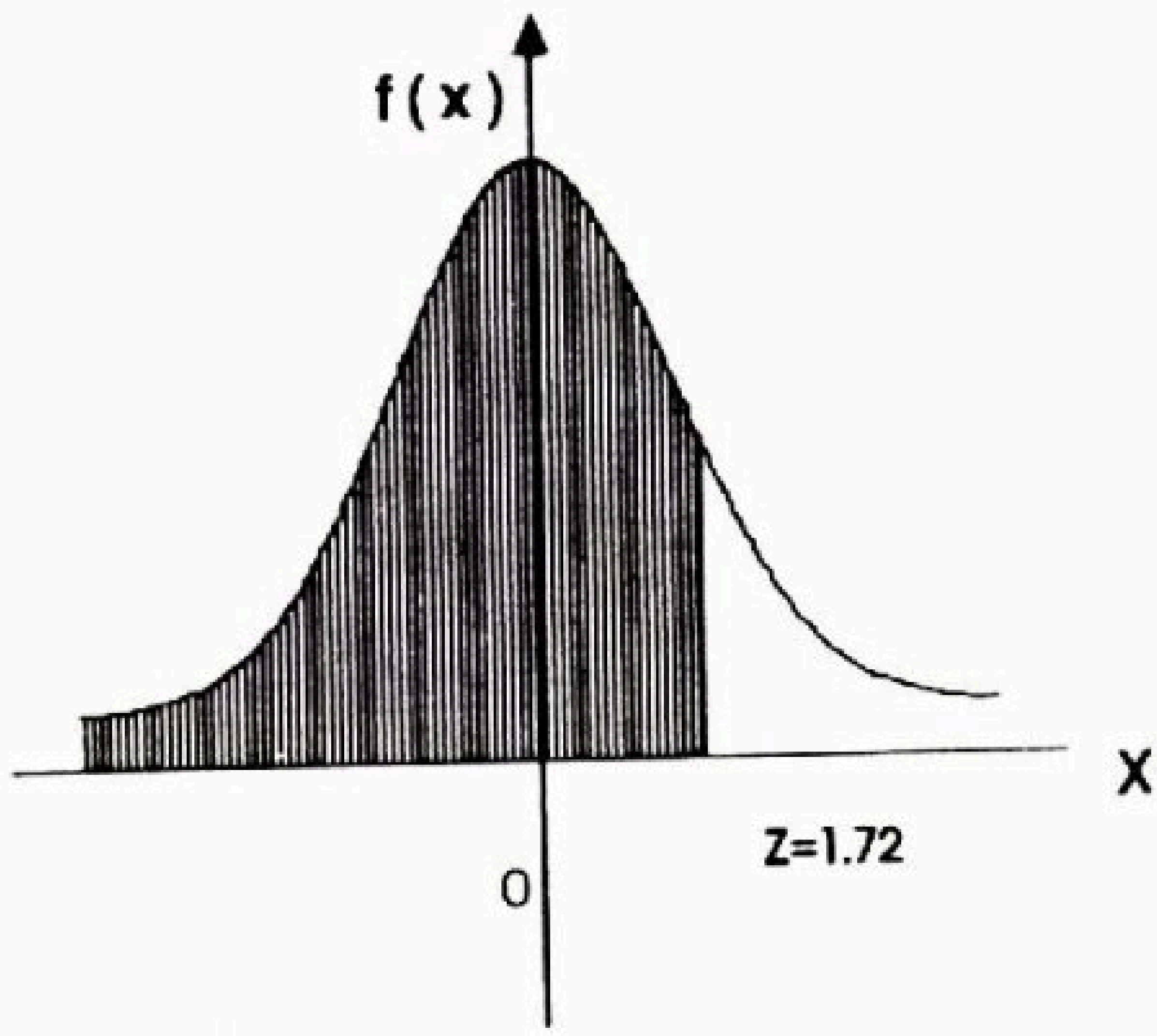
الحـــل

حيث إن المتغيّر العشوائي Z موزع توزيعًا قياسيًا

(i)

$$P(Z \leq 1.72) = \int_{-\infty}^{1.72} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2} z}\, dz$$

أي أن قيمة الاحتمال عبارة عن المساحة المظللة في شكل (٩-١٣) التالي:



شكل (٩ - ١٣) يمثل Φ (1.72)

ولقد قام الإحصائيون بعمل جداول لحساب المساحات تحت المنحنى الطبيعي القياسي في الفترة من 3 - إلى 3 + وذلك لأن 99.75% من هذه القيم تقع داخل هذه الفترة والعمود الأول من الجدول يحتوي على القيم:

$$-0.0\ ,\ -0.1\ ,\ -0.2\ ,\ \ldots\ ,\ -3.0\ ,\ 0.0\ ,\ 0.1\ ,\ \ldots\ ,\ 3$$

والأعمدة التسعة الباقية ذات عناوين

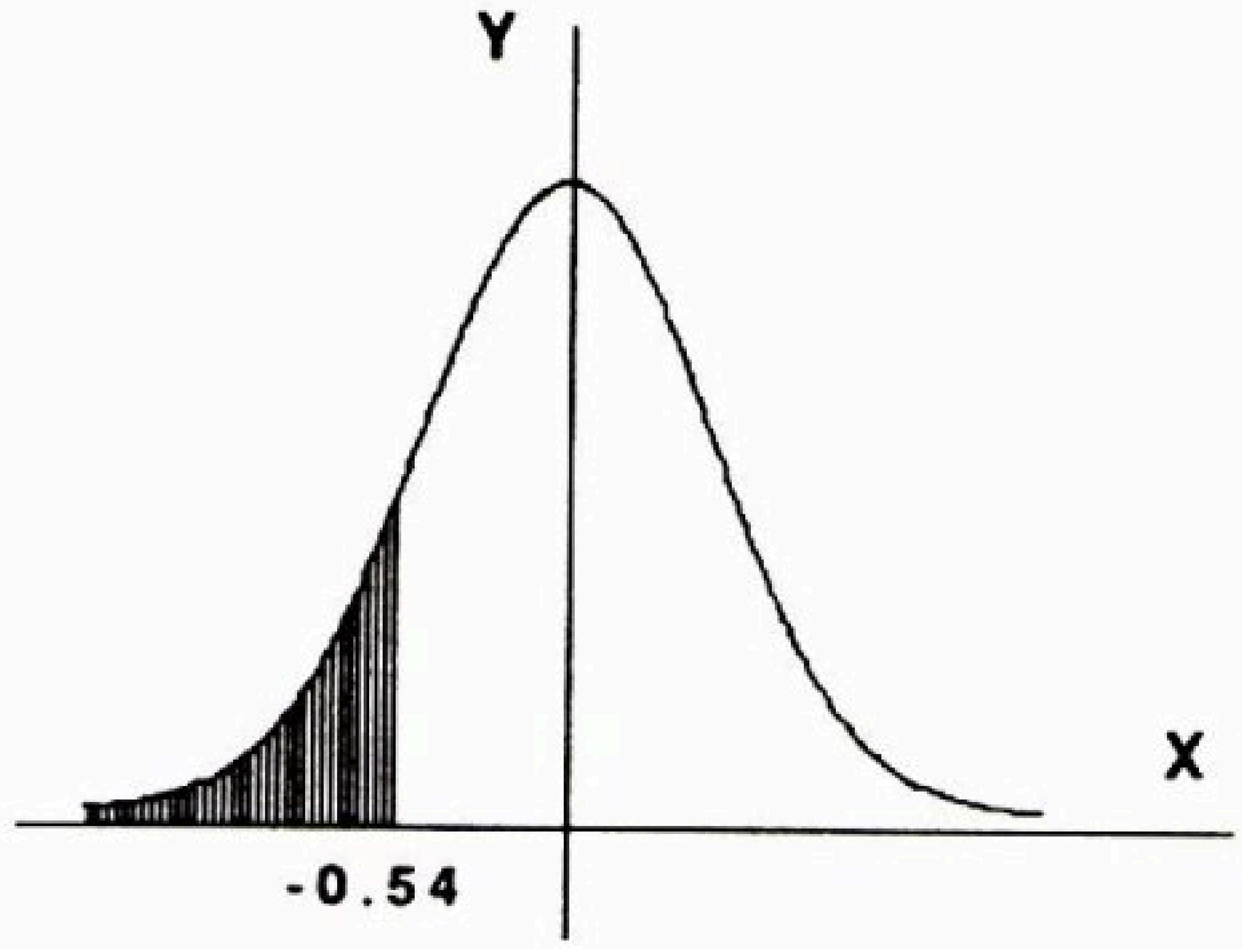
$$0.00\ ,\ 0.01\ ,\ 0.02\ ,\ \ldots\ldots\ldots\ ,\ 0.09$$

وبذلك لإيجاد $P(Z < 1.72)$ نبحث في العمود الأول عن القيمة 1.7 ثم نتحرك أفقيًا في الصف نفسه يمينًا حتى نصل إلى القراءة التي أسفل الرقم 0.02 فتكون هي القيمة المطلوبة أي أن:

$$\Phi(1.72) = P(Z \leqslant 1.72) = 0.9573$$

(ii) $$P(Z \leqslant -0.54)$$

هذا الاحتمال يساوي المساحة المظللة للمنحنى الطبيعي القياسي التالي: شكل (٩ ـ ١٤).



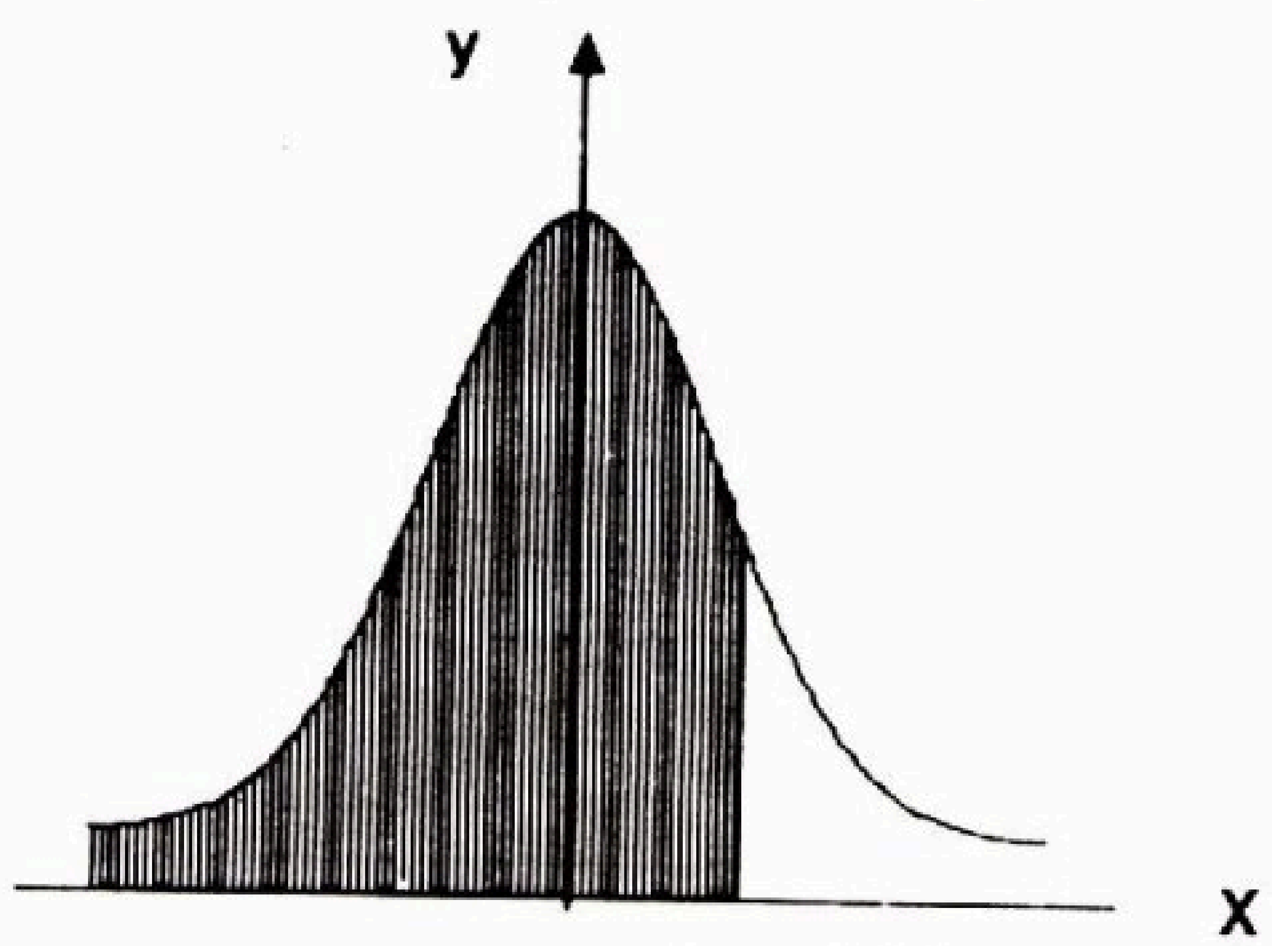
شكل (٩ ـ ١٤) الجزء المظلّل يمثل $\Phi(-0.54)$

نبحث في العمـود الأول من الجـدول عن القيمة 0.5 - ثم نتحرك أفقيًا حتى نحصل على القراءة التي تحت القيمة 0.04 فتكون هي الاحتمال المطلوب

$$P(Z \leqslant -0.54) = 0.2946$$

(iii) $$P(Z \leqslant 1.07)$$

وهذا الاحتمال يساوي الجزء المظلل في الشكل (٩ ـ ١٥):



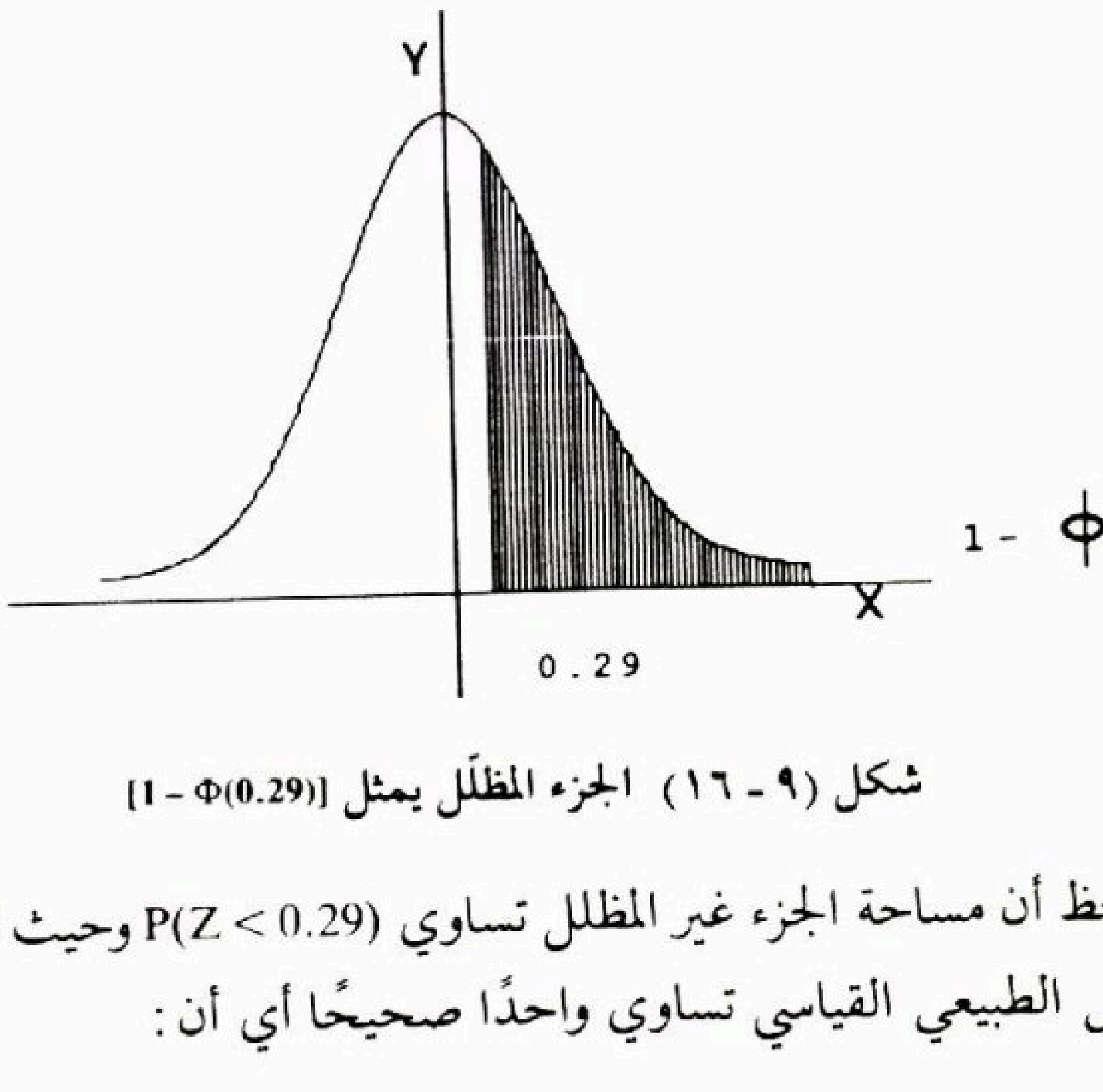
شكل (٩ ـ ١٥) الجزء المظلّل يمثل $\Phi(1.07)$

ومن الجدول نحصل على

$$P(Z \leqslant 1.07) = 0.8577$$

(iv) $$P(Z \geqslant 0.29)$$

هذا الاحتمال يساوي الجزء المظلل في شكل (٩ ـ ١٦) التالي :



شكل (٩ ـ ١٦) الجزء المظلّل يمثل [1 - Φ(0.29)]

نلاحظ أن مساحة الجزء غير المظلل تساوي $P(Z < 0.29)$ وحيث المساحة الكلية تحت المنحنى الطبيعي القياسي تساوي واحدًا صحيحًا أي أن :

$$P(Z \geqslant 0.29) + P(Z < 0.29) = 1$$

$$P(Z \geqslant 0.29) = 1 - P(Z < 0.29)$$

$$= 1 - \Phi(0.29)$$

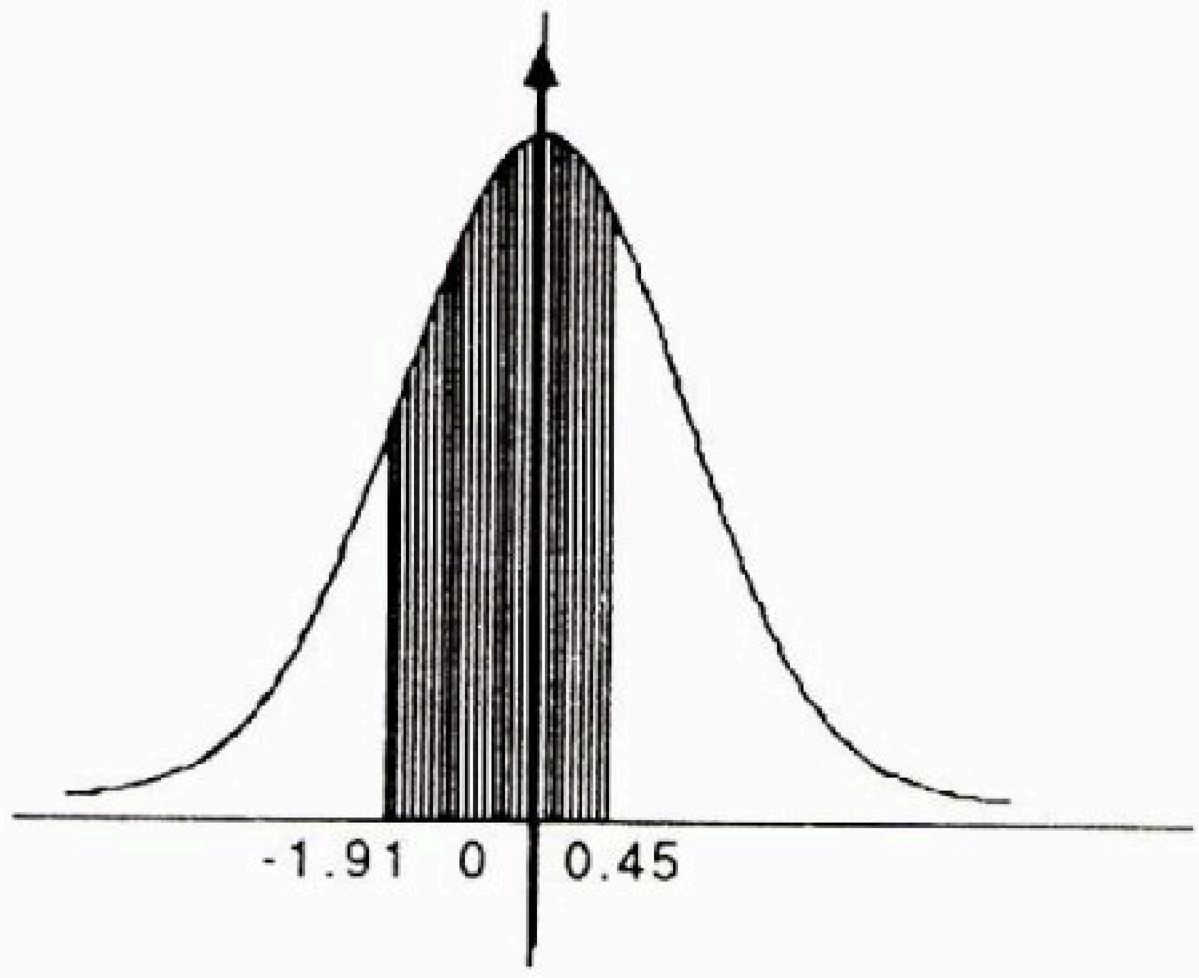
ومن الجداول الخاصة بالمنحنى الطبيعي القياسي نحصل على :

$$P(Z \geqslant 0.29) = 1 - 0.6141$$

$$= 0.3859$$

(v) $$P(-1.91 \leqslant Z \leqslant 0.45)$$ (١)

وهذا الاحتمال يساوي الجزء المظلل في شكل (٩ ـ ١٧) التالي:

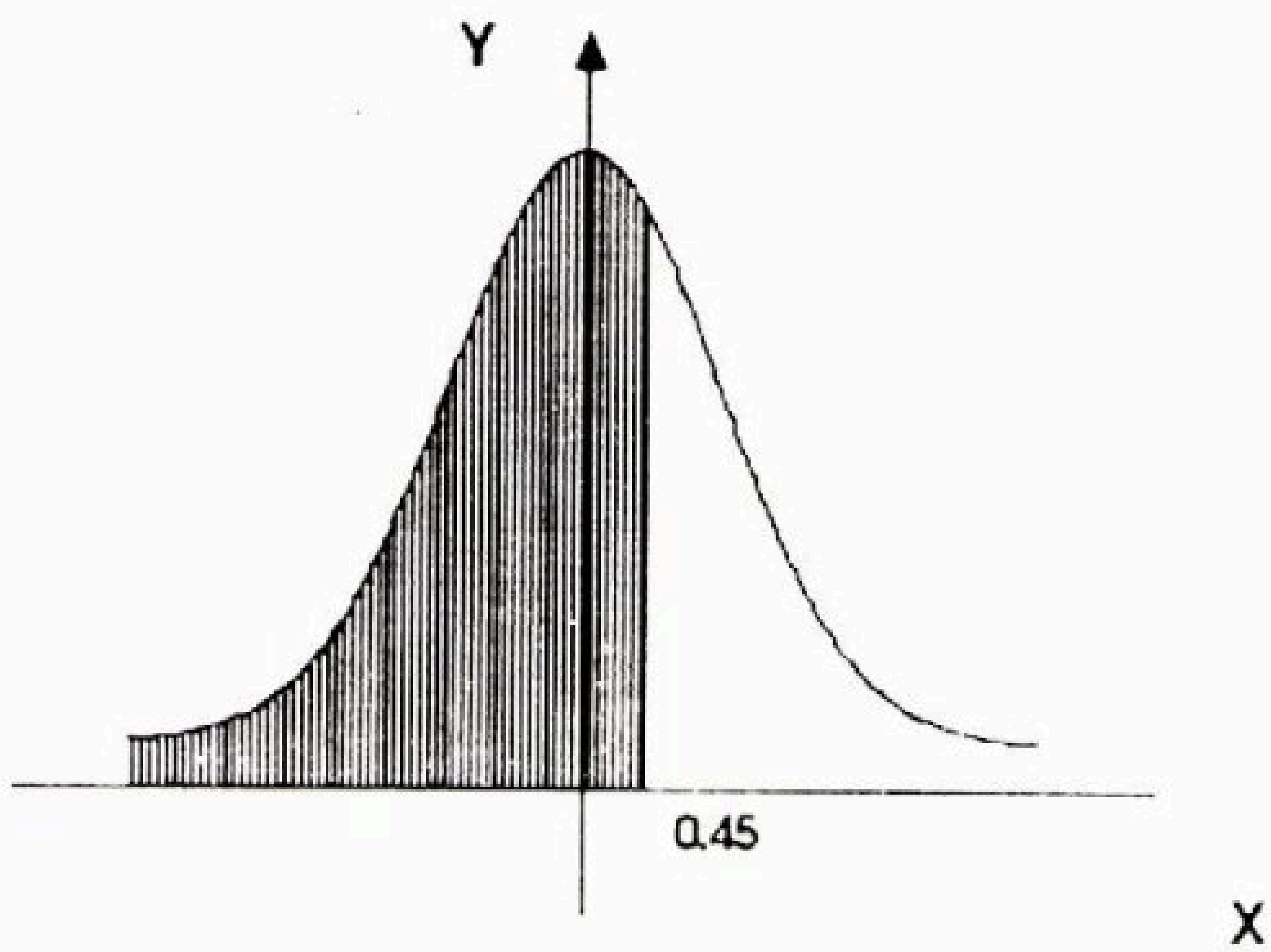


شكل (٩ ـ ١٧) الجزء المظلّل يمثل [$\Phi(0.45) - \Phi(-1.91)$]

ونلاحظ أن الشكل (٩ ـ ١٧) السابق يمكن تجزئته كالتالي:

$$P(Z \leqslant 0.45) = 0.6736 \qquad (٢)$$

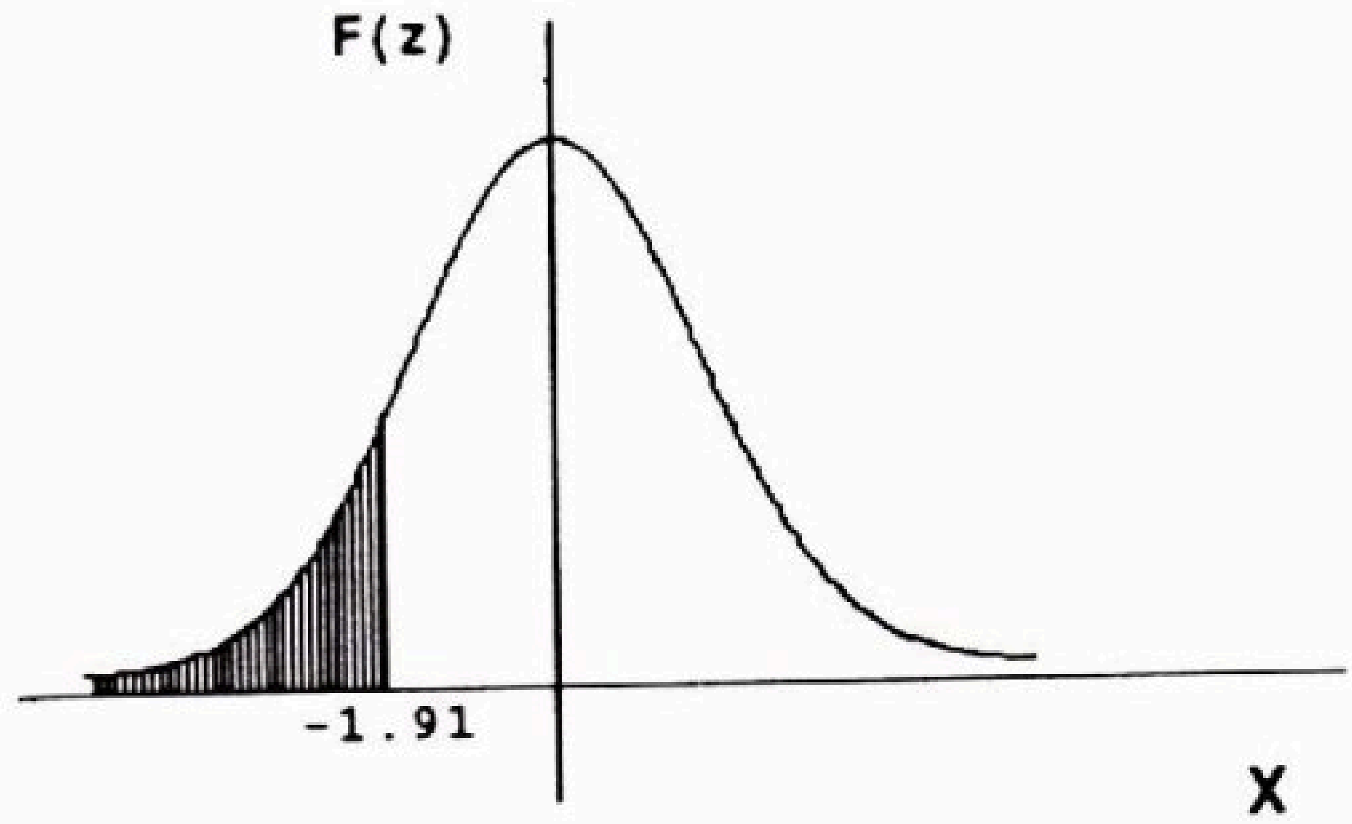
وتساوي المساحة الظلّلة في الشكل التالي شكل (٩ ـ ١٨):



شكل (٩ ـ ١٨) الجزء المظلّل يمثل $\Phi(0.45)$

(٣) $$P(Z \leqslant -1.91) = 0.0281$$

وتساوي المساحة المظلّلة في الشكل التالي (شكل ٩ ـ ١٩):



شكل (٩ ـ ١٩) الجزء المظلّل يمثل $\Phi(-1.91)$

المساحة في (١) المظلّلة = المساحة (٢) المظللة - المساحة (٣) المظلّلة في الأشكال الثلاثة السابقة وتحسب كالتالي:

$$\begin{aligned} P(-1.91 \leqslant Z \leqslant 0.45) &= P(Z \leqslant 0.45) - P(Z \leqslant -1.91) \\ &= 0.6736 - 0.0281 \\ &= 0.6455 \end{aligned}$$

(٩ ـ ١٠ ـ ٣) إيجاد الاحتمالات لبعض قيم المتغيّر العشوائي الطبيعي

لأي قيمة محدّدة x للمتغيّر العشوائي الطبيعي $X \sim N(\mu, \sigma^2)$ توجد قيمة مقابلة لها $Z = \frac{x-\mu}{\sigma}$ وتسمى بالقيمة القياسية وذلك لأن متوسط المتغيّر العشوائي الذي يمثلها يساوي صفرًا وتباينة يساوي واحدًا صحيحًا

$$P(X \leqslant x) = P\left(\frac{X-\mu}{\sigma} \leqslant \frac{x-\mu}{\sigma}\right)$$

$$P\left(Z \leqslant \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

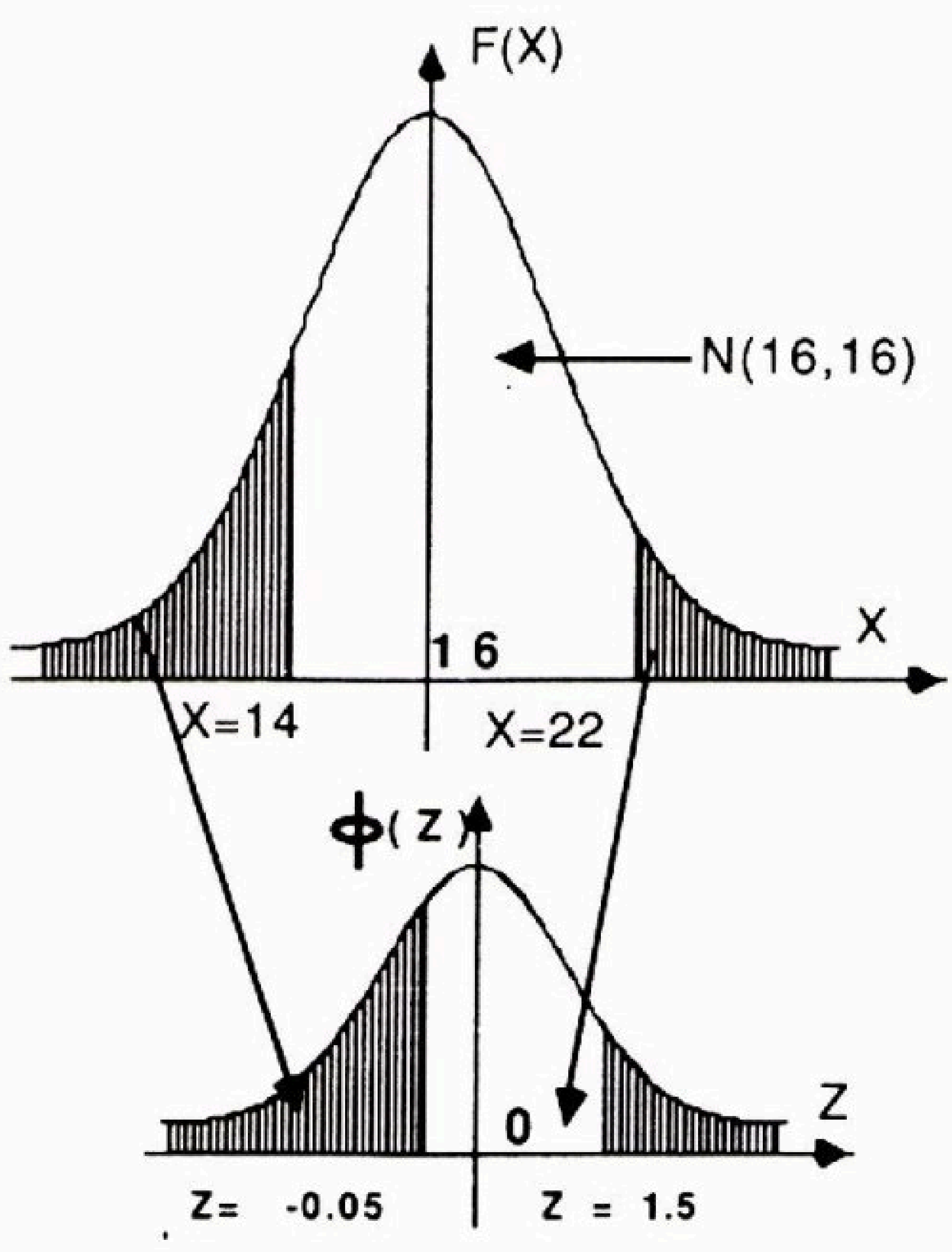
مثال (٩ ـ ٢٧)

إذا كان $X \sim N(16, 16)$ فأوجد :

$$\text{(a)}\ P(X \leq 14) \qquad \text{(b)}\ P(X \geq 22)$$

الحـــل

يمكن تمثيل المسـاحة المظلّلة المطلوبة في منحنى التوزيع الطبيعي $X \sim N(16, 16)$ والمساحة المظلّلة المناظرة للتوزيع الطبيعي القياسي $Z \sim N(0, 1)$ كما هو موضح بالرسم التالي في شكل (٩ ـ ٢٠) :



شكل (٩ ـ ٢٠) يمثل تقابل المساحات بين التوزيع الطبيعي N(16, 16) والتوزيع الطبيعي القياسي N(0, 1)

$$(a)\ \ x = 14 \Leftarrow z = \frac{14-16}{4} = \frac{-2}{4} = -0.5$$

$$P(X \leqslant 14) = P(Z \leqslant -0.5) = 0.3085$$

$$(b)\ \ x = 22 \Leftarrow z = \frac{22-16}{4} = \frac{6}{4} = 1.5$$

$$P(X \geqslant 22) = P(Z \geqslant 1.5) = 1 - P(Z < 1.5)$$

$$= 1 - 0.9332 = 0.0668$$

ويمكن استخدام الصيغ الاحتمالية التالية لحساب الاحتمالات

$$(a)\quad P(X \leqslant a) = P(Z \leqslant \frac{a-\mu}{\sigma})$$

$$(b)\quad P(X \geqslant a) = P(Z \geqslant \frac{a-\mu}{\sigma}) = 1 - P(Z \leqslant \frac{a-\mu}{\sigma})$$

$$(c)\quad P(a \leqslant X \leqslant b) = P(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma})$$

$$= P(Z \leqslant \frac{b-\mu}{\sigma}) - P(Z \leqslant \frac{a-\mu}{\sigma})$$

$$= \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$$

حيث a ، b قيمتان ثابتتان للمتغير العشوائي $X \sim N(\mu, \sigma)$.

مثال (٩ ـ ٢٨)

فترة الحمل التامة في الإنسان تعتبر متغيّر عشوائي طبيعي بمتوسط 266 يوم وانحراف معياري 12 يومًا. ما هي نسبة السيدات الحوامل اللاتي يستمر حملهن بين 260 و 270 يوم.

الحل

نفرض أن المتغيّر العشوائي X يمثل فترة الحمل التامة للإنسان أي أن:
$X \sim N(266,144)$ وبذلك يكون المطلوب $P(260 \leqslant X \leqslant 270)$

وباستخدام الصيغة (C) السابقة نحصل على التالي:

$$P(260 \leqslant X \leqslant 270) = P(Z \leqslant \frac{270-266}{12}) - P(Z \leqslant \frac{260-266}{12})$$

$$= P(Z \leqslant 0.33) - P(Z \leqslant -0.5)$$

$$= 0.6293 - 0.3085$$

$$= 0.3208$$

أي أن حوالي 32% من السيدات الحوامل يستمر حملهن ما بين 260 و 270 يومًا.

التوزيع N(266, 144) هو عبارة عن نموذج مثالي لتصرف أو لتغير فترة الحمل لدى الإنسان واستخدام هذا النموذج بطريقة غير صحيحة قد يؤدي إلى نتائج منافية للعقل مثلاً:

$$P(X \geqslant 314) = P(Z \geqslant \frac{314-266}{12})$$

$$= P(Z > 4)$$

$$= 1 - P(Z < 4)$$

$$= 1 - 0.99997 = 0.00008$$

وهكذا فإن احتمال هذا الحدث ليس صفرًا بالرغم من أن $X \geqslant 314$ يكون حدثًا مستحيلاً لأن 314 يومًا تساوي 10.5 شهرًا، ولأن كل حالات الحمل تنتهي بالتأكيد قبل 10.5 شهرًا.

(٩ ـ ١٠ ـ ٤) التقريب الطبيعي لتوزيع ذي الحدين

يستخدم التوزيع الطبيعي (وهو توزيع متصل) لتقريب توزيع ذي الحدين (وهو توزيع متقطّع) وذلك عندما تكون n عدد المحاولات كبيرًا ($n > 20$) و P احتمال النجاح قريب من 0.5 أو إذا كان $[np > 5, np(1-p) > 5]$.

مثال (٩ ـ ٢٩)

عند رمي قطعة عملة متزنة 10 مرات وكان المتغير العشوائي X يمثل عدد الصور. فأوجد $P(X = 2)$.

الحــل

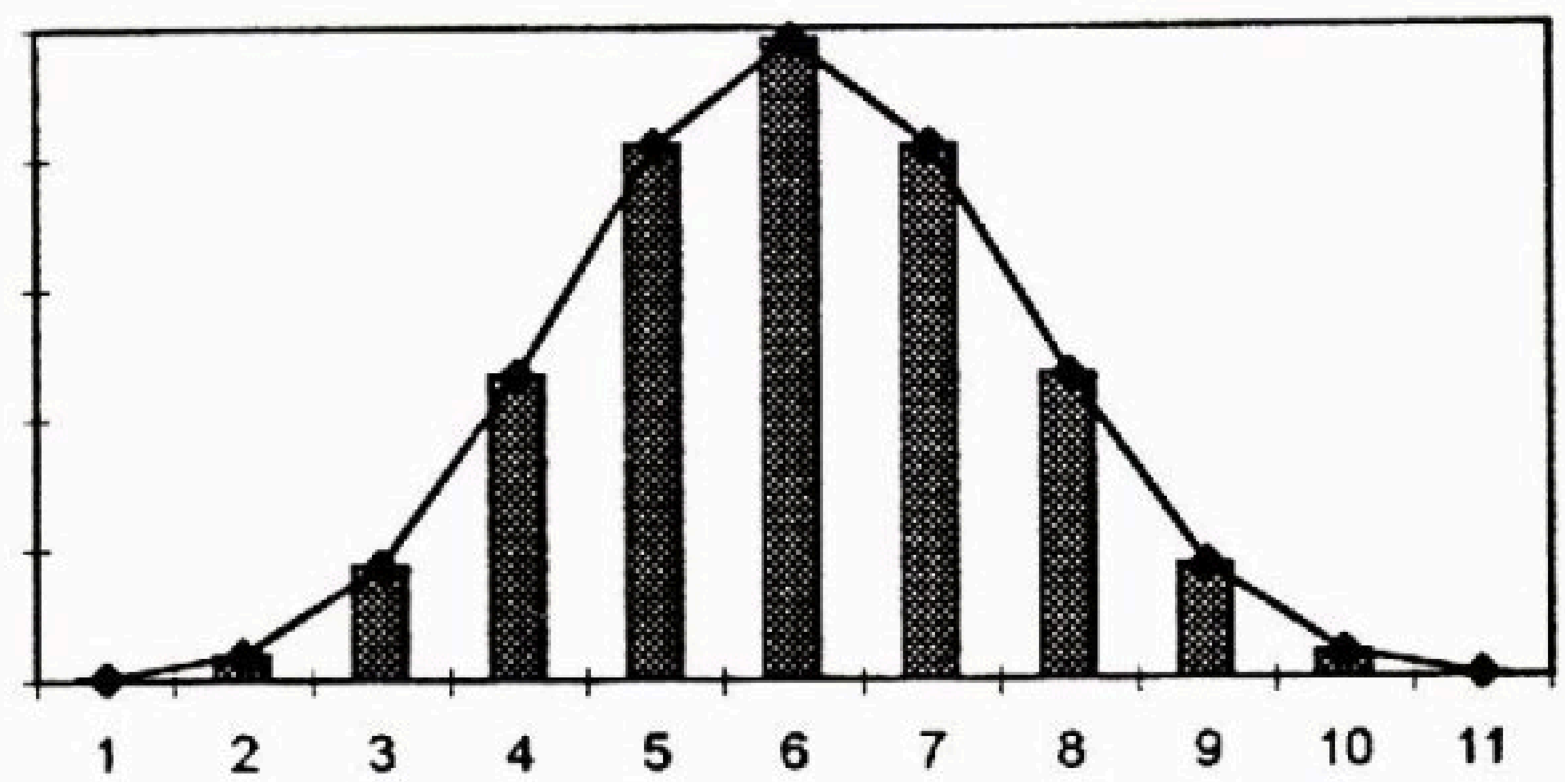
المتغيّر العشوائي X يتبع توزيع ذي الحدين.

$$f(x\,;n,p) = \binom{n}{x} p^x q^{n-x}\ ,\ x = 0, 1, \dots , n$$

بالمعالم

$$P = \frac{1}{2}\quad ,\quad q = \frac{1}{2}\quad ,\quad n = 10$$

يكون $P(X = 2)$ هو الجزء المظلّل (المستطيل) الموضح في الشكل التالي: (شكل ٩ ـ ٢١).



شكل (٩ ـ ٢١) يمثل المدرج التكراري لرمي قطعة عملة عشرة مرات

وهـذه المسـاحة تساوي $\binom{10}{2}(\frac{1}{2})^2(\frac{1}{2})^8$ وتساوي 0.04394 ولكن المساحة السـابقـة يمكن تقـريبها بالمساحة تحت المنحنى من 1.5 إلى 2.5 أي يمكن التقريب بالتوزيع الطبيعي أي أن:

$$P(X = 2) \simeq P(1.5 \leq X \leq 2.5)$$

ويمكن حساب μ , σ للتوزيع ذي الحدين كالتالي:

$$\mu = np = 10\left(\frac{1}{2}\right) = 5\ ,\ \sigma = \sqrt{npq} = \sqrt{10\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 1.58$$

$$P(1.5 \leqslant X \leqslant 2.5) = P(\frac{1.5-5}{1.58} \leqslant Z \leqslant \frac{2.5-5}{1.58})$$

$$= P(-2.22 \leqslant Z \leqslant -1.58)$$

$$= P(Z \leqslant -1.58) - P(Z \leqslant -2.22)$$

$$= 0.05705 - 0.01321$$

$$= 0.04384$$

ونلاحظ أن التقريب جيد (بالرغم من صغر n).

نظرية (٩ ـ ٣)

إن كان X متغيراً عشوائيًا له توزيع ذي حدين بمتوسط $\mu = np$ وانحراف معياري $\sigma = \sqrt{np(1-p)}$ وكانت قيمة p قريبة من $\frac{1}{2}$ وكانت قيمة n كبيرة. بحيث $np > 5$, $np(1-p) > 5$ معًا فإن:

$$P(X = a) \simeq P(a - \frac{1}{2} \leqslant X \leqslant a + \frac{1}{2})$$

$$= P(\frac{a - np - 0.5}{\sqrt{np(1-p)}} \leqslant Z \leqslant \frac{a - np + 0.5}{\sqrt{np\,(1-P)}})$$

حيث Z هو المتغيّر العشوائي الطبيعي القياسي.

مثال (٩ ـ ٣٠)

ما هو احتمال الحصول على 6 صور في 16 رمية لقطعة عملة متزنة. باستخدام توزيع ذي الحدين ثم بالتقريب الطبيعي لذي الحدين.

الحـــل

الاحتمال المطلوب هو $P(X = 6)$

وبتطبيق التوزيع ذي الحدين

$$P(X=6) = \binom{16}{6}\left(\frac{1}{2}\right)^6\left(1-\frac{1}{2}\right)^{16-6} = \frac{8008}{65536} = 0.1222$$

الآن نوجد هذا الاحتمال مستخدمين النظرية السابقة نحسب μ و σ كالتالي:

$$\mu = np = 16 \times \frac{1}{2} = 8$$

$$\sigma = \sqrt{npq} = \sqrt{16 \times \frac{1}{2} \times \frac{1}{2}} = 2$$

$$P(X=6) \{ P\left(\frac{6-8-0.5}{2} \leqslant Z \leqslant \frac{6-8+0.5}{2}\right)$$

$$= P(-1.25 \leqslant Z \leqslant -0.75)$$

$$= \Phi(-0.75) - \Phi(-1.25)$$

$$= 0.2266 - 0.1056$$

$$= 0.1210$$

وهذا تقريب جيد.

مثال (٩ - ٣١)

ما هو احتمال الحصول على 70 حالة نجاح على الأقل في 100 محاولة إن كان احتمال النجاح في المحاولة الواحدة 0.75 ؟

الحـــل

الاحتمال المطلوب هو: $P(X \geqslant 70)$

وأن

$$P(X \geqslant 70) = p(X=70) + P(X=71) + \ldots + P(X=100)$$

$$= \binom{100}{70}(0.75)^{70}(0.25)^{30} + \binom{100}{71}(0.75)^{71}(0.75)^{29} +$$

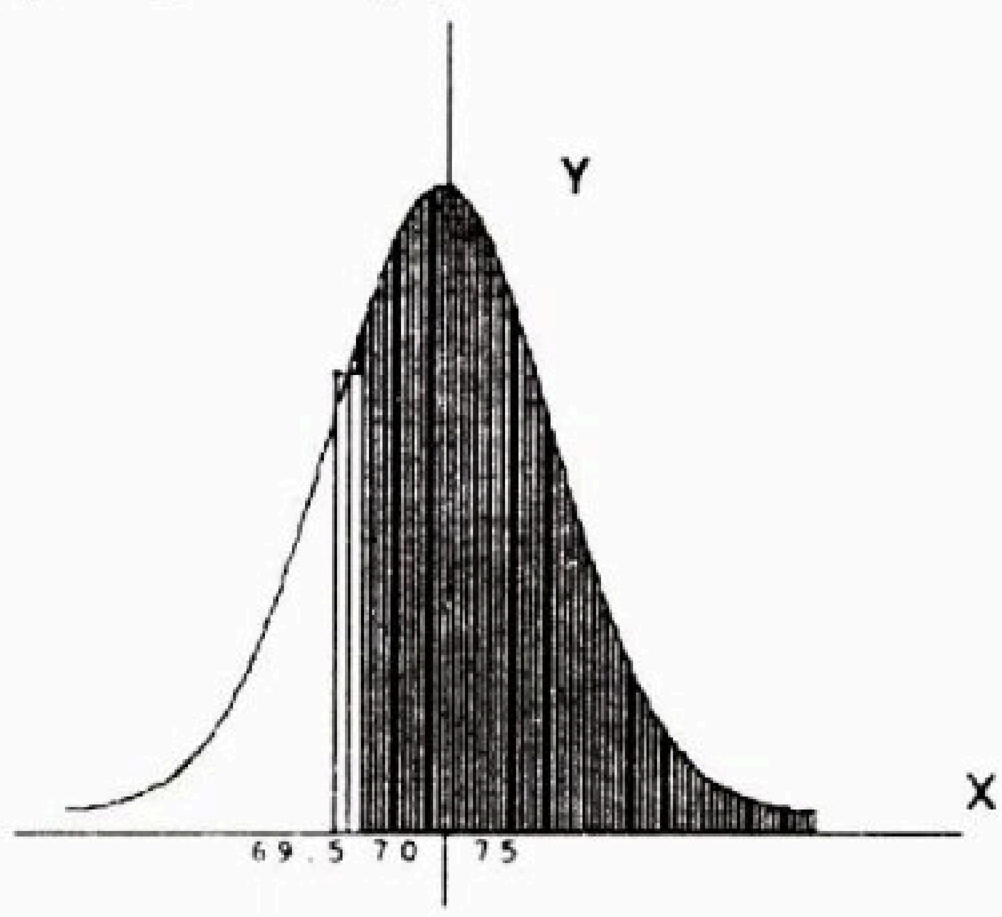
$$+ \ldots\ldots + \binom{100}{100}(0.75)^{100}(0.25)^{0} = 0.892$$

وكما نرى أن إيجاد هذا يحتاج إلى الكثير من الوقت والكثير من الحسابات ولكن يمكن إيجاده بسهولة بالتقريب الطبيعي .

$$\mu = np = 100\,(0.75) = 75\ ,\ \sigma = \sqrt{npq} = \sqrt{100(0.75)\,(0.25)} = 4.33$$

$$\begin{aligned} P(X \geqslant 70) &= P(X \geqslant 69.5) \\ &= P(Z \geqslant \frac{69.5 - 75}{4.33}) \\ &= P(Z \geqslant -1.2702) \\ &= 1 - P(Z \leqslant -1.27) \\ &= 1 - 0.1020 = 0.8980 \end{aligned}$$

ونوضح المساحة المطلوبة بالجزء المظلل في الشكل التالي:شكل (٩ ـ ٢٢) .



شكل (٩ ـ ٢٢) يبين الاحتمال المطلوب في مثال (٩ ـ ٣١) باستخدام تقريب التوزيع الطبيعي

(٩ ـ ١١) تمـــارين

(١) أي من القيم التـالية يمكن أن تكون قيمًا لتوزيع احتمالي (دالة كتلة احتمالية) للمتغيّر العشوائي X والذي يأخذ القيم 1,2,3,4 ، وعلل ذلك .

a) $f(1) = 0.26$; $f(2) = 0.26$; $f(3) = 0.26$; $f(4) = 0.26$

b) $f(1) = \frac{1}{9}$; $f(2) = \frac{2}{9}$; $f(3) = \frac{1}{3}$; $f(4) = \frac{1}{3}$

c) $f(1) = 0.15$; $f(2) = 0.28$; $f(3) = 0.29$; $f(4) = 0.28$

d) $f(1) = 0.33$; $f(2) = 0.37$; $f(3) = -0.3$; $f(4) = 0.33$

e) $f(1) = \frac{1}{4}$; $f(2) = \frac{1}{8}$; $f(3) = \frac{1}{16}$; $f(4) = \frac{1}{32}$

(٢) أوجد قيم الثابت c في الدوال التالية والتي تجعل هذه الدوال دوال كتلة احتمالية :

a) $f(x) = cx$, حيث $x = 1, 2, 3, 4, 5$

b) $f(x) = c\binom{5}{x}$, حيث $x = 0, 1, 2, 3, 4, 5$

(٣) إذا كان X متغيرًا عشوائيًا دالة كتلته الاحتمالية معطاة بالجدول التالي :

x	−1	2	5	8
f(x)	0.3	a	0.2	0.1

i) احسب قيمة الثابت a .

ii) احسب توقع وتباين المتغيّر العشوائي X .

iii) احسب $P(X \geq 0)$, $P(0 \leq X \leq 5)$, $P(X > 9)$, $P(X \leq -2)$

(٤) حدد كون كل من الدالّتين التاليتين تصلح كدالة كتلة احتمالية للمتغيّر العشوائي X .

i) $f(x) = \frac{1}{7}$; $x = 0,1,2,3,4$

ii)

x	0	1	2	3	4
f(x)	0.1	−0.1	0.5	0.3	0.2

(٥) عند دخولك إلى الجامعة تواجه إشارتين ضوئيتين تعملان مستقلتين بعضهما عن بعض واحتمال كلتيهما أن تكون حمراء عند وصولك إليها هو 0.5 . لنرمز بالرمز R للإشارة الحمراء وبالرمز G للإشارة الخضراء.

i) اكتب فضاء العيّنة واحسب الاحتمالات المرافقة لكل نقطة عيّنة.

ii) إذا كان المتغيّر العشوائي X يمثل عدد الإشارات الحمراء التي تواجهها فما هي قيمة X عند كل نقطة عيّنة ثم أكتب دالة الكتلة الاحتمالية للمتغيّر العشوائي X .

(٦) متغيّر عشوائي X يأخذ القيم الآتية {1, 2, 3, 4} . بيّن كون الدالة $f(x) = \frac{x^2}{30}$ تصلح دالة توزيع احتمالي مع التعليق ومن ثم أوجد:

i) توقع وتباين المتغيّر العشوائي X .

ii) احتمال أن تكون X تساوي ثلاثة على الأكثر.

iii) احتمال أن تكون X تساوي 2 على الأقل.

(٧) ألقي حجرا نرد مرة واحدة فكان المتغيّر العشوائي X يمثل مجموع الرقمين اللذين يظهران إلى أعلى. أوجد قيم المتغيّر العشوائي X وكذلك دالة الكتلة الاحتمالية $f_X(x)$ ثم أرسمها وكذلك التوقع والتباين لهذا المتغيّر العشوائي.

(٨) صنعت قطة نقود بحيث كان $P(H) = \frac{3}{4}$ و $P(T) = \frac{1}{4}$ ألقيت هذه القطعة 4 مرات فإن كان المتغيّر العشوائي X يمثل عدد الصور. أوجد دالة الكتلة الاحتمالية $f_X(x)$ وكذلك المتوسط μ والتباين σ^2 .

(٩) إن كان X متغيراً عشوائيًّا يتبع التوزيع ذي الحدّين فأثبت أن:

i) $\mu_X = np$

ii) $\sigma_X^2 = n\,p\,q$

(١٠) إذا علم أن المتغير العشوائي X يتبع التوزيع فوق الهندسي فأثبت أن:

i) $\mu_X = np$

ii) $\sigma_X^2 = n\,p\,q\,(\frac{N-n}{N-1})$

(١١) إن علم أن المتغير العشوائي X له التوزيع فوق الهندسي ودالة كتلته الاحتمالية هي:

$$f^X(x) = \frac{\binom{NP}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$$

فأثبت أن دالة الكتلة الاحتمالية لهذا المتغير X تؤول إلى دالة الكتلة الاحتمالية لتوزيع ذي الحدين التالي:

$$P_X(x) = \binom{n}{x} P^x q^{n-x}$$

وذلك عندما N تكون كبيرة جدًّا.

(١٢) لوحظ في إحدى الألعاب الرياضية التي نتيجتها إما فوز أو خسارة أن احتمال فوز لاعب ما ثابت في أي مباراة ويساوي 0.6 فإن علم أن هذا اللاعب سوف يلعب 5 مباريات مع أشخاص مختلفين خلال الموسم القادم وكان المتغيّر العشوائي X يمثل عدد مرات الفوز.

أوجد:

i) عدد المباريات المتوقع أن يفوز بها اللاعب.

ii) أحسب الانحراف المعياري للمتغيّر العشوائي X .

iii) احتمال أن يفوز بأربع مباريات على الأكثر.

iv) احتمال أن يخسر مباراتين على الأكثر.

(١٣) يقوم قسم علم الحيوان بتربية نوع نادر من الأسماك مع نوع آخر مقارب للنّوع النادر في الحجم وذلك لدراسة حياة هذه الأنواع معًا. فإذا احتوت البركة الصناعية في القسم على 15 سمكة منها سبعة من النوع النادر فإن طلب أستاذ المعمل من المساعد اختيار 4 سمكات عشوائيًا ووضعها في صندوق زجاجي مملوء بالماء وكان المتغيّر العشوائي X يمثل عدد السمكات من النوع النادر في العينة المختارة والمطلوب:

i) إيجاد دالة الكتلة الاحتمالية للمتغيّر العشوائي X .

ii) حساب التوقع الرياضي للمتغيّر العشوائي X .

iii) حساب تباين المتغيّر العشوائي X .

(١٤) وعاء يحتوي على 4 كرات مرقمة بالأرقام 4 ,3 ,2 ,1 على التوالي سحبت كرتان من الوعاء بدون إرجاع. عرّف المتغيّر العشوائي X على أنه مجموع ما يظهر على الكرتين المسحوبتين.

i) أوجد القيم الممكنة للمتغيّر العشوائي X والاحتمال المناظر لهذه القيم. ثم أوجد متوسط وتباين المتغيّر العشوائي X .

ii) أحسب المطلوب في (i) إن كان السحب بإرجاع.

(١٥) سحبت كرتان على التوالي بدون إرجاع من وعاء يحتوي على 4 كرات حمراء و 3 كرات سوداء، فإن كان المتغيّر العشوائي X يمثل عدد الكرات الحمراء في العيّنة المسحوبة. أوجد دالة التوزيع الاحتمالي (دالة الكتلة الاحتمالية) للمتغيّر العشوائي X . وإن كان السحب بإرجاع، أحسب دالة الكتلة الاحتمالية في هذه الحالة أيضًا.

(١٦) باقة من الزهور فيها 12 زهرة بيضاء، 4 زهور حمراء اخترنا عشوائيًّا مع الإعادة (بإرجاع) ثلاث زهرات للتأكد من رائحتها. وليكن X عدد الزهور الحمراء التي نحصل عليها.

i) أكتب جدول التوزيع الاحتمالي للمتغيّر العشوائي X .

ii) ما هو احتمال الحصول على زهرة واحدة حمراء على الأقل.

iii) أحسب توقع وتباين المتغيّر العشوائي X .

(١٧) إن كان احتمال إصابة قنّاص لهدف هو 0.3 . فإن صوب نحو الهدف 5 مرات متتالية، وإن عرّفنا المتغيّر العشوائي X بأنه عدد مرات الإصابة.

i) أوجد دالة الكتلة الاحتمالية للمتغيّر العشوائي X .

ii) أوجد توقع وتباين المتغيّر العشوائي X .

iii) أوجد احتمال أن يصيب الشخص الهدف مرة واحدة على الأكثر.

(١٨) لدينا 10 صمامات كهربائية منها 5 لا تعمل. اخترنا عشوائيًّا عيّنة من ثلاث صمامات. وليكن X عدد الصمامات التي لا تعمل في العيّنة.

i) أكتب دالة الكتلة الاحتمالية للمتغيّر العشوائي X .

ii) أحسب أن تكون العيّنة كلها سليمة.

iii) أحسب احتمال أن تتضمن العيّنة صمامًا واحدًا لا يعمل على الأقل.

(١٩) اختبار متعدد الاختيارات مكون من 6 أسئلة كل سؤال له 3 إجابات واحدة فقط منها صحيحة. إن أجاب أحد الطلبة بالطريقة التالية:

رمي زهرة نرد متزنة، ثم يختار الجواب الأول إن ظهر له 1 أو 2 . ويختار الجواب الثاني إن ظهر له 3 أو 4 . ويختار الجواب الثالث إن ظهر له 5 أو 6 .

ما هو احتمال أن يجيب الطالب على:

i) ثلاث إجابات صحيحة.

ii) ولا إجابة صحيحة.

iii) على الأكثر خمس إجابات صحيحة .

iv) أوجد متوسط وتباين التوزيع أعلاه .

(٢٠) ادّعى مهندس الأمن والسلامة أن حادثة واحدة من كل عشر حوادث للمرور تعزى إلى إرهاق السائق .

أوجد احتمال أن 3 من بين 5 حوادث سير تنسب لذلك السبب . ثم أوجد متوسط وتباين الحوادث التي تعزى إلى إرهاق السائق .

(٢١) إن كان من بين 16 متنافسًا لوظيفة ما عشرة لهم درجات جامعية اختير 3 متنافسين عشوائيًا للمعاينة .

أوجد الاحتمالات التالية :

i) لا يوجد بينهم من يحمل درجة جامعية .

ii) واحد فقط يحمل درجة جامعية .

iii) اثنان يحملان درجة جامعية .

iv) المتنافسون الثلاثة يحملون درجات جامعية .

(٢٢) شحنة من 80 جهازًا كهربائيًا من بينها 4 أجهزة متعطلة . اختيرت 3 أجهزة عشوائيًا .

أوجد احتمال أن تحتوي هذه العيّنة على جهاز واحد متعطل .

(٢٣) صندوق يتضمن ثماني تفاحات اثنتان منها تالفة سحبنا عشوائيًا ثلاثًا منها وكان المتغيّر العشوائي X يمثل عدد التفاحات التالفة التي حصلنا عليها .

أكتب دالة الكتلة الاحتمالية للمتغيّر العشوائي X في الحالتين التاليتين :

i) السحب مع الإعادة .

ii) السحب بدون إعادة .

(٢٤) إن كانت نسبة المعيب في الإنتاج تمثل % 10 سحبت عيّنة مكونة من 5 وحدات .
فأوجد الاحتمالات التالية :
i) لا يوجد في العيّنة وحدات معيبة .
ii) توجد وحدة واحدة معيبة .
iii) يوجد على الأكثر وحدة معيبة .
iv) يوجد على الأقل وحدتان معيبتان .

(٢٥) إن كان احتمال أن يتخرج طالب من الجامعة هو $\frac{3}{5}$. سحبت عيّنة مكونة من 4 طلاب .
أوجد الاحتمالات التالية :
i) أن تخرج جميع الطلاب في العينة .
ii) أن يتخرج طالبان فقط .
iii) أن يتخرج طالبان على الأقل .

(٢٦) مصنع به 15 عاملاً و 5 من المهندسين سحبت عينة عشوائية مكونة من ثلاثة أفراد .
أوجد الاحتمالات التالية :
i) العيّنة كلها من المهندسين .
ii) العيّنة بها عامل واحد ومهندسان .
iii) العيّنة كلها من العمال .

(٢٧) إن كانت نتيجة النجاح في الامتحان لشعبة مكونة من 40 طالبًا هي % 80 سحب عشوائيًا عيّنة من 5 طلاب .
أحسب الاحتمالات :
i) نجاح ثلاثة طلّاب من بينهم .
ii) على الأقل طالب واحد ناجح .
iii) على الأكثر طالب واحد ناجح .

(٢٨) إن كان عدد البراكين في السنة متغيرًا عشوائيًا يتبع تقريبًا التوزيع الطبيعي بمتوسط $\mu = 20.8$ وانحراف معياري $\sigma = 4.5$.

أوجد الاحتمالات التالية:

i) أن يحدث 18 بركانًا بالضبط في السنة.

ii) أن يحدث على الأقل 22 بركانًا في السنة.

iii) أن يحدث ما بين 20 إلى 30 بركانًا في السنة.

(٢٩) إن كان 70 % من الأشخاص المسافرين عبر المحيط الإطلنطي يشعرون بدوار البحر، فما هو احتمال أن من بين 150 مسافرًا عبر المحيط الإطلنطي على الأقل 100 شخص يشعرون بدوار البحر.

(٣٠) عدد شهادات الزواج التي تصدر في مدينة ما في أحد الشهور كان لها متوسط $\mu = 134$ ، وانحراف معياري $\sigma = 7.5$.

أجب عما يأتي:

i) ماذا تقول نظرية شيبشيف بـ $k = 9$ عن عدد حالات الزواج في ذلك الشهر؟

ii) وفقًا لهذه النظرية ما هو احتمال أن ما بين 74 و 194 شهادة زواج تصدر خلال ذلك الشهر على الأقل؟

(٣١) إن كان توزيع بواسون يعطي بالدالة $f(x, \lambda)$ كالتالي:

$$f(X, \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \quad , \quad \lambda > 0 \ , \ x = 0, 1, 2, \ldots$$

أوجد:

$$f(3, \frac{1}{5}) \ , \ f(5,2) \ , \ f(6,1)$$

(٣٢) إن كانت نسبة المصابين بمرض ما في بلد معيّن هو 0.003 فما هو احتمال عدم وجود أي إصابة في حي يسكنه 6000 نسمة؟

(٣٣) إن كان هنـاك 300 خطأ مطبعي موزعـة على صفحات كتاب به 600 صفحة اختيرت صفحة بطريقة عشوائية.

أوجد احتمال أن تحتوي هذه الصفحة على:

i) على الأكثر خطأ واحد مطبعي.

ii) تحتوي على ثلاثة أخطاء فقط.

(٣٤) إن كان المتغيّر العشوائي Z له توزيع طبيعي قياسي،

فأوجد الاحتمالات التالية:

(i	$\Phi(0.32)$	$P(Z<1.8)$
(ii	$\Phi(1.25)$	$P(Z>-0.5)$
(iii	$\Phi(-0.82)$	$P(-0.2<Z<0.5)$

(٣٥) إن كانت أطوال 500 ورقة من أوراق نبات معين لها توزيع طبيعي بمتوسط 132 مليمترًا وانحراف معياري 10 مليمترًا.

أوجد عدد الأوراق التالية:

i) ما بين 130 ملم، 140 ملم.

ii) أكبر من 150 ملم.

iii) أقل من 130 ملم.

(٣٦) إن كانت درجات أحد الامتحانات لمجموعة من الطلاّب يتبع التوزيع الطبيعي بالقيم $\mu = 74, \sigma = 12$.

أوجد الدرجات بالوحدات القياسية للطلبة الحاصلين على:

i) 65 , ii) 74 , iii) 86 , vi) 92 درجة

(٣٧) من بيانات تمرين 36 أوجد الدرجات المناظرة للدرجات القياسية التالية:

i) -1 , ii) 0.5 , iii) 1.25 , iv) 1.7

(٣٨) أوجد المساحة تحت منحنى التوزيع الطبيعي القياسي التي تقع :

i) بين $Z = 0$, $Z = 0.87$ ، ii) بين $Z = -1.66$, $Z = 0$

iii) على يمين $Z = 0.48$ ، iv) على يسار $Z = 1.3$

v) على يسار $Z = -0.79$ ، vi) بين $Z = 0.55$ و $Z = 1.12$

vii) بين $Z = -1.05$ و $Z = -1.75$

(٣٩) إن كان متـوسط المتغيّر العشوائي X الذي يتبع التوزيع الطبيعي يساوي 80 وانحرافه المعياري يساوي 4.8 .

أوجد الاحتمالات التي يأخذها المتغيّر العشوائي للقيم التالية :

i) أقل من 87.2 ، ii) أكبر من 76.4

iii) بين 81.2 و 86.0 ، iv) بين 71.6 و 88.4

(٤٠) إن كانت درجة الحرارة خلال شهر مارس تتبع التوزيع الطبيعي بأحد البلاد بتوقع °20 وانحراف معياري °3.33 .

أوجد احتمال أن تكون درجة الحرارة بين °21.11 , °26.66 في هذا الشهر.

(٤١) إن كان متوسط المتغيّر العشوائي الموزع توزيعًا طبيعيًا هو 62.4 .

أوجد الانحراف المعياري له إن علم أن %20 من المساحة تحت المنحنى تقع على يمين 79.2 .

(٤٢) إن كان الانحراف المعياري σ للمتغيّر العشوائي الذي يتبع التوزيع الطبيعي يساوي 5 ، أوجد متوسطه إن علم أنه يأخذ قيمة أقل من 52.5 باحتمال يساوي 0.8264 .

(٤٣) إن كان زمن الاحتراق بالنسبة لصاروخ تجريبي متغيراً عشوائيًا موزعًا توزيعًا

طبيعيًا . حيث متوسطه يساوي 4.3 ثانية وانحرافه المعياري يساوي 0.04 ثانية .
أوجد الاحتمالات التالية :

i) أن يحترق هذا الصاروخ في أقل من 4.25 ثانية .

ii) أن يحترق هذا الصاروخ في أكثر من 4.40 ثانية .

iii) أن يحترق هذا الصاروخ فيما بين 4.30 و 4.42 ثانية .

(٤٤) i) إن كان المتغيّر العشوائي X له توزيع $N(7, \sigma^2)$ وكان 20% من مساحة منحنى الكثافة الاحتمالية يقع على يمين العدد 9 فأحسب σ .

ii) إن كان X متغيراً عشوائيًا له توزيع $N(\mu, 100)$ أحسب μ إن كان احتمال أن يأخذ قيمة أقل 80.5 هو 0.3264 .

(٤٥) إن كانت القيمة $Z_{\alpha/2}$ هي القيمة للمتغير العشوائي Z التي تقع على يمينها مساحة تساوي $\frac{\alpha}{2}$ مع ملاحظة أن :

$$P(-z_{a/2} < Z < z_{a/2}) = 1 - \alpha$$

حدد قيم $z_{\alpha/2}$ عندما α تأخذ القيم التالية .

$\alpha = 0.05$ (iii, $\alpha = 0.1$ (ii, $\alpha = 0.01$ (i

(٤٦) وجدنا أن الفترة الزمنية الضرورية لإنجاز اختبار للذكاء يخص طلبة إحدى الكليات يتوزع احتماليًا وفق التوزيع الطبيعي بمتوسط 70 دقيقة وانحراف معياري يساوي 12 دقيقة .

كم يجب أن نحدد زمن الاختبار إن أردنا إتاحة وقت كافٍ لـ 90% من الطلّاب لإتمام الاختبار.

(٤٧) إن كان في إنتاج إحدى الآلات 30% معيبًا . أخذنا عيّنة من 120 قطعة .

فأوجد باستخدام التقريب الطبيعي لذي الحدين الاحتمالات التالية :

i) أن يكون 30 وحدة معيبة فقط .

i) أن يكون 40 وحدة معيبة على الأكثر.

iii) أن يكون 50 وحدة معيبة على الأقل.

(٤٨) إن قذفنا قطعة عملة 80 مرة وكانت العملة متزنة.

أوجد الاحتمالات التالية:

i) الحصول على 25 صورة فقط.

ii) على الأقل الحصول على 30 صورة.

iii) على الأكثر الحصول على 45 صورة.

(٤٩) إن كان 60% من السحاب يظهر نموًا بأيونات الفضة.

أوجد احتمال من بين 60 سحابة 30 سحابة تظهر نموًا على الأكثر.

(٥٠) إن كانت قيم المتغيّر العشوائي X تختلف عن متوسط المجتمع بالمقادير التالية:

10 (iii) 9 (ii) 8 (i)

فأوجد احتمال انحراف قيم المتغيّر العشوائي عن المتوسط في الحالات السابقة إن كان الانحراف المعياري لهذا المجتمع $\sigma = 4$.

(٥١) إن كانت $k = 3$. أوجد أقل احتمال انحراف قيم المتغيّر العشوائي X عن متوسطه.

(٥٢) إن كان مقدار اختلاف متغيّر عشوائي عن متوسطه أقل أو يساوي 4.5 إن كانت $k = 1.5$.

أوجد الانحراف المعياري له باستخدام نظرية شيبشيف ثم أوجد أقل احتمال لمقدار اختلاف المتغيّر العشوائي عن متوسطه.

(٥٣) إن علم أن درجات طلّاب السنة الأولى بكلية العلوم هو متغيّر عشوائي X يتبع

التوزيع الطبيعي بمتوسط $(\mu=67)$ وتباين $(\sigma^2=64)$. اختير طالب بشكل عشوائي .

i) ما احتمال أن تكون درجته بين 75 و 65 .

ii) إن كان عدد الطلاّب المسجلين بكلية العلوم للسنة الأولى يساوي 600 طالب أوجد عدد الطلاّب الذين تزيد درجاتهم عن 60 .

(٥٤) إن كان احتمال أن يتأخر طالب عن موعد الامتحان هو 0.1

i) ما هو احتمال أن يتأخر طالبان عن امتحان مقرر من بين عشرة طلاّب مسجلين .

ii) ما هو احتمال أن يتأخر 25 طالبًا على الأقل عن امتحان مقرر بين 350 طالبًا مسجلين في هذا المقرر.

(٥٥) إن كانت درجة الحرارة خلال فترة من العام في بلد ما تتبع التوزيع الطبيعي بمتوسط $(\mu = 20°)$ وانحراف معياري $(\sigma = 3°)$.

أوجد الاحتمالات التالية :

i) أن لا تزيد درجة الحرارة عن °23 .

ii) أن تكون درجة الحرارة بين °26 و °15 .

iii) أن لا تقل درجة الحرارة عن °20 .

iv) ما هي درجة الحرارة التي تتجاوزها الحرارة في البلد باحتمال مقداره 0.937 .

(٥٦) i) المتغيّر العشوائي Z يتبع التوزيع الطبيعي القياسي (المعياري) ما هو احتمال أن يأخذ Z قيمًا.

(i) أكبر من 1.24 . (ii) أقل من 0.46 . (iii) ما بين 0.36 و 0.23 −

ii) نعلم من سجلات سابقة أن 25% من جميع المرضى الذين يتناولون دواء معيّنًا تظهر عليهم أعراض جانبية من بين مائة شخص استخدموا هذا الدواء. ما هو احتمال أن يعاني 26 منهم على الأقل أعراضًا جانبية .

(٥٧) إن كانت نسبة الرحلات القادمة من جدة إلى الصالة الداخلية في مطار الملك خالد الدولي بالرياض هي 25% من مجموع الرحلات الداخلية الواصلة إلى المطار من مختلف أنحاء المملكة.

i) أوجد احتمال أن توجد طائرتان قادمتان من جدة من بين 4 طائرات وصلت إلى الصالة الداخلية بالمطار.

ii) إن كان عدد الرحلات الداخلية التي وصلت إلى المطار في أحد الأيام هو 30 رحلة أوجد الاحتمال التقريبي (مستخدمًا التوزيع الطبيعي) في أن يكون 5 منها على الأقل قادمة من جدة.

(٥٨) لديك صندوق به 5 كرات حمراء، 7 كرات بيضاء، سحبت ثلاث كرات بدون إرجاع.

أحسب الاحتمالات التالية:

i) أن تكون الكرة الأولى حمراء.

ii) أن تكون الكرة الثانية حمراء.

iii) أن تكون الكرة الثالثة حمراء.

ماذا يمكن أن نستنتج؟

(٥٩) ضع إشارة (✓) إذا كان الجواب صحيحًا وإلا إشارة (X) إذا كان الجواب خطأ:

ا - القيمة المتوقعة E(x) لمتغير عشوائي X هي عمليًا متوسط القيم التي يمكن أن يفترضها المتغير العشوائي X على المدى الطويل ☐

ب - في التوزيع ذي الحدين يتغير احتمال النجاح من تكرار إلى آخر ☐

جـ - التوزيع الطبيعي هو توزيع متناظر (متماثل) ☐

د - المساحة تحت منحنى دالة كثافة احتمالية يجب أن يساوي واحد ☐

(٦٠) أكمل كل من الفراغات التالية:

ا - المتغير العشوائي هو دالة حقيقية معرفة على

ب ـ كي تصلح دالة f(x) كدالة كثافة احتمالية لمتغير عشوائي مستمر يجب أن تحقق الشرطين:

(i) .... (ii) ....

جـ ـ في التوزيع الثنائي يبقى .... ثابتًا من محاولة (تكرار) إلى أخرى.

د ـ إذا كان n عدد المحاولات (التكرارات) و p احتمال النجاح في توزيع ثنائي (ذي الحدين) فإن متوسط هذا التوزيع .... = E(x) وتباين هذا التوزيع .... = V(x).

١٠

المعاينة وتوزيعاتها

● مقدمة ● المجتمعات المنتهية وغير المنتهية
● المعاينة العشوائية ● توزيعات المعاينة
● نظرية النهاية المركزية ● تمارين

(١٠ - ١) مقدمة

سبق لنا دراسة بعض التوزيعات المهمة في الفصل التاسع مثل التوزيع ذي الحدين، وفوق الهندسي، وتوزيع بواسون للمتغيرات العشوائية المنفصلة، والتوزيع الطبيعي للمتغيّرات العشوائية المتصلة. ولكن ما زال أمامنا تساؤل وهو كيف نستخدم هذه التوزيعات في التطبيقات الإحصائية؟

لا شك أن هناك ميادين كثيرة تستخدم فيها هذه التوزيعات، منها طرق أخذ العيّنات التي توفر كثيراً من الجهد والوقت والمال، بدلاً من استخدام البيانات من التعداد الشامل، لحساب المعالم المجهولة مثل المتوسط والتباين للمجتمع محل الدراسة، وتحسب من العيّنة إحصائيات لتلك المعالم المجهولة للمجتمع.

فمثلاً إذا أخذنا عيّنة حجمها (n) من طلّاب جامعة الملك سعود الذي حجمه (N) وكان المطلوب تقدير متوسط الطول للطلّاب باستخدام متوسط العيّنة $\bar{x}$. فالسؤال هنا ما هو توزيع الإحصائية $\bar{x}$ عندما نأخذ جميع العيّنات ذات الحجم المتساوي n من هذا المجتمع الطلّابي؟ وما هي خصائص هذا التوزيع مثل التوقع والتباين؟ مثل هذا

التوزيع هو ما يسمى بتوزيع المعاينة، وسوف نتناوله في هذا الفصل بالشرح والتفصيل والأمثلة كما يلي.

(١٠ - ٢) المجتمعات المنتهية وغير المنتهية

المجتمـع

يتكـون من جميع المشـاهـدات الممكنة (ولو نظريًّا) لظاهرة معطاة مثل ظاهرة الطول، والوزن للإنسان... إلخ.

العيّنـة

هي جزء من المشاهدات الممكنة لظاهرة معيّنة مثل أطوال شعبة معيّنة من طلّاب جامعة الملك سعود... إلخ.

المجتمع المنتهِ

ويتكون من عدد محدود أو منتهٍ من المشاهدات مثل عدد طلّاب جامعة الملك سعود في عام دراسي معيّن.

المجتمع المنتهي

وهـو يتكون من عدد غير محدود من المشاهدات مثل عدد الأسماك في الخليج العربي، عدد الطيور المهاجرة في فصل الخريف من الشمال... إلخ.

(١٠ - ٣) المعاينة العشوائية

عدد العينات المختلفة ذات الحجم n من مجتمع منتهٍ ذي حجم N عندما يكون السحب بدون إرجـاع هو $\binom{N}{n}$ عيّنة. ومثال على ذلك يوجد عدد $\binom{12}{2}$ يساوي 66 عيّنة مختلفة ذات حجم 2 من مجتمع يتكون من 12 عنصرًا. أما إذا كان السحب بإرجاع فإن عدد العيّنات المختلفة هو N^n أي يوجد 12^2 يساوي 144 عيّنة مختلفة ذات حجم 2 من مجتمع يحتوي على 12 عنصرًا.

(١٠ - ٣ - ١) العيّنة العشوائية البسيطة

هي العيّنة التي تختار من مجتمع محدود أو منتهٍ بحيث يكون لكل عيّنة من $\binom{N}{n}$

من العيّنـات المختلفة الاحتمال $1/\binom{N}{n}$ نفسه في الاختيار عندما يكون السحب بدون إرجاع ويكون لكل عيّنة من N^n من العيّنات المختلفة الاحتمال $\frac{1}{N^n}$ نفسه في الاختيار.

مثال (١٠ ـ ١)

أوجد جميع العيّنات الممكنة التي حجمها 3 والمسحوبة بدون إرجاع من المجتمع $\{a, b, c, d, e\}$ ثم بين كيف يمكن اختيار عيّنة عشوائية من هذه العيّنات.

الحـل

لأن العيّنات الممكنة هو $\binom{5}{3}$ ويساوي 10 وهذه العيّنات هي:

abc, abd, abe, acd, ace, adc, bcd, bce, bde, cde

فإذا اخـترنـا واحـدة من هذه الـعـيّنـات بحـيث يكـون احتـمال اختـيار أي منها يساوي $\frac{1}{10}$ فإننا نسمي هذه العيّنة (عيّنة عشوائية)

ولاختيار عيّنة بطريقة عشوائية هناك عدة طرق نذكر منها:

١) نكتب كل العيّنات الممكنة والتي عددها $\binom{N}{n}$ على بطاقات، ثم نخلطها جيدًا ثم نسحب منها واحدة بدون النظر إليها. ولكن هذه الطريقة غير مناسبة عندما تكون كل من n ، N كبيرتين.

٢) تكتب جميع المشاهدات التي عددها N على بطاقات ثم نختار عيّنة عشوائية حجمها n بحيث يكون احتمال اختيار أي بطاقة من البطاقات متساوٍ.

٣) استخدام الأعداد العشوائية لاختيار العيّنة العشوائية. ونوضح ذلك كما يلي:-

فإذا كانت مدينة ما تحتوي على 247 صيدلية، وأردنا اختيار 12 صيدلية منها بطريقة عشوائية أي إيجاد عيّنة عشوائية ذات حجم 12 . فإننا نرقم الصيدليات كالتالي:

001 , 002 , ... , 246 , 247

(مثلاً ترقيم الصيدليات حسب ترتيبها في دليل التليفونات أو أي طريقة أخرى).

ثم نستخدم جداول الأعداد العشوائية (جدول 2) آخر الكتاب فتكون

إحدى العيّنات العشوائية هي الصيدليات ذات الأرقام التالية :

136 , 216 , 184 , 018 , 041 , 153 , 215 , 037 , 174 , 189 , 024 , 044

ملاحظة

عنـد اختيار هذه الأرقـام اخـترنـا أي ثلاثة أعمدة رأسية من جداول الأرقام العشـوائية لأن المجتمـع الأصلي محل الدراسة مكوَّن من ثلاث خانات وأهملنا جميع الأرقام في الجدول التي هي أكبر من 247 كما أهملنا الأرقام المكررة .

(١٠ ـ ٤) توزيعات المعاينة

Sampling distribution

(١٠ ـ ٤ ـ ١) التوزيع العيني النظري للمتوسط $\bar{X}$

ولتوضيح توزيع المعاينة للمتوسط $\bar{X}$ اعتبر المثال التالي :

مثال(١٠ ـ ٢)

مجتمع مكوّن من المفردات :

3, 5, 7, 9, 11

والمطلـوب

i) حساب متوسط هذا المجتمع μ' وانحرافه المعياري σ .

ii) أكتب جميع العيّنات الممكنة ذات حجم $n = 2$ في حالة السحب بدون إرجاع .

iii) أوجد دالّة الكتلة الاحتمالية للمتغيّر العشوائي $\bar{X}$ $(f(\bar{x}))$ ثم أرسمها .

ix) المتغير العشوائي $\bar{X}$ أوجد توقعه $\mu_{\bar{x}}$ وتباينه $\sigma^2_{\bar{x}}$.

الحـــل

(i
$$\mu = \frac{3+5+7+9+11}{5} = 7$$

$$\sigma = \sqrt{\frac{(3-7)^2 + (5-7)^2 + (7-7)^2 + (9-7)^2 + (11-7)^2}{5}}$$

$$= \sqrt{8} = 2.828$$

ii) وحيث إن السحب بدون إرجاع فإن العيّنات الممكنة هي $\binom{5}{2}$ وتساوي 10 وهي كالتالي:

$$\{3,5\}, \{3,7\}, \{3,9\}, \{3,11\}, \{5,7\}, \{5,9\}, \{5,11\}, \{7,9\}$$
$$\{7,11\}, \{9,11\},$$

iii)

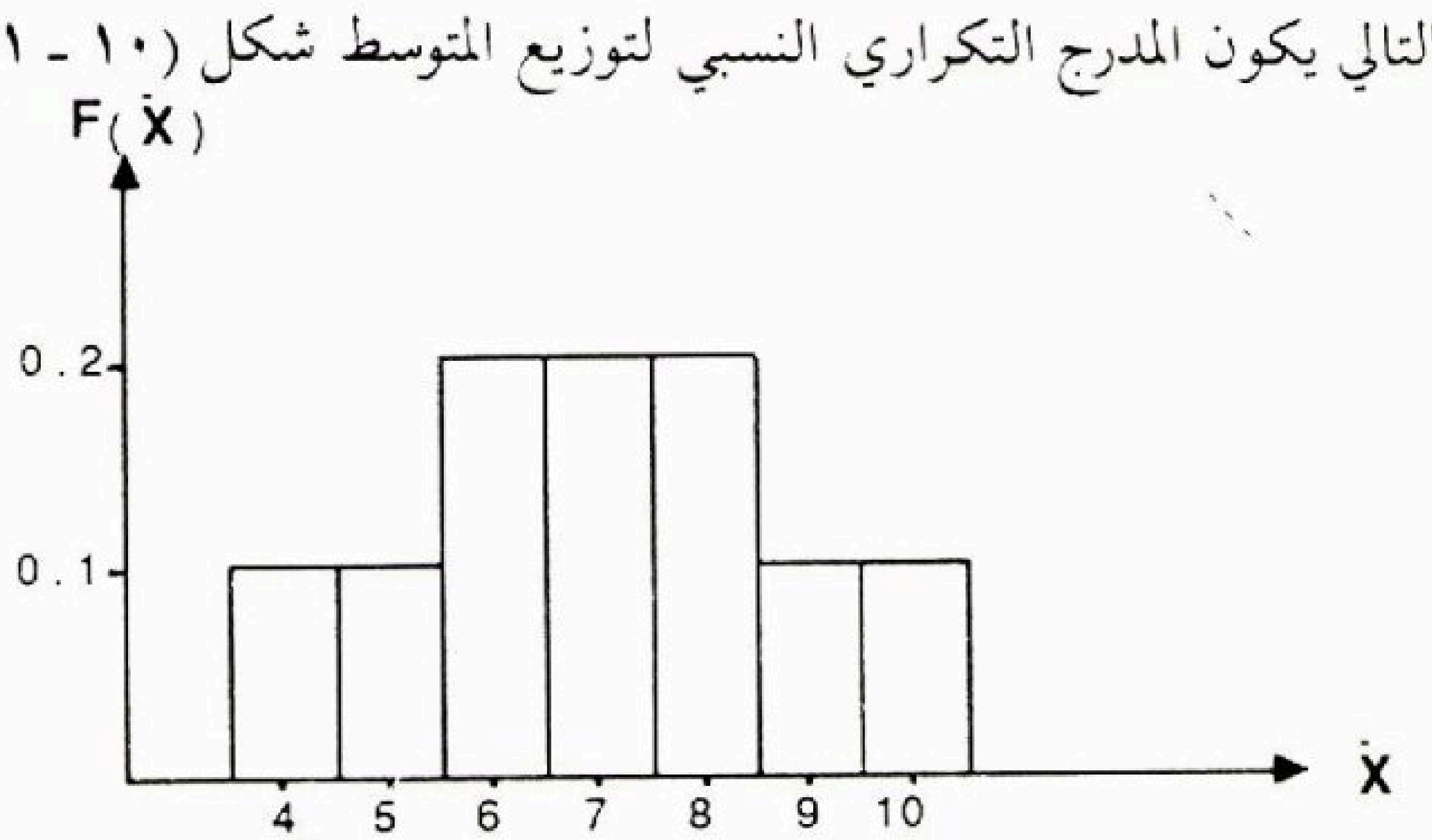
ومتوسطات هذه العيّنات هي على الترتيب كالتالي:

4, 5, 6, 7, 6, 7, 8, 8, 9, 10

وإذا كانت المعاينة عشوائية فإن كل عيّنة من العيّنات السابقة يكون لها احتمال قدره $\frac{1}{10}$. وبذلك يمكن تلخيص التوزيع العيني للمتوسط $\bar{X}$ أي $(f(\bar{x}))$ بالجدول التالي:

$\bar{x}$	4	5	6	7	8	9	10
$f(\bar{x})$	**0.1**	**0.1**	**0.2**	**0.2**	**0.2**	**0.1**	**0.1**

وبالتالي يكون المدرج التكراري النسبي لتوزيع المتوسط شكل (١٠ ـ ١) التالي:



شكل (١٠ ـ ١) يمثل توزيع المعاينة للمتوسط $\bar{x}$

ونلاحظ من الجدول التكراري السابق ما يلي :

١ - الاحتمال = $\frac{6}{10}$ بأن قيم $\bar{x}$ لا تختلف عن μ بأكثر أو بأقل من 1 .

٢ - الاحتمال = $\frac{8}{10}$ بأن قيم $\bar{x}$ لا تختلف عن μ بأكثر أو بأقل من 2 .

(د) توقع $\bar{X}$ ويرمز له $\mu_{\bar{x}}$ أو $E(\bar{X})$ ويحسب كالتالي :

$$\mu_{\bar{X}} = \sum_{\bar{x}} \bar{x}\, f(\bar{x})$$

$$= 4\times0.1 + 5\times0.1 + 6\times0.2 + 7\times0.2 + 8\times0.2 + 9\times0.1 + 10\times0.1$$

$$\therefore\ \mu_{\bar{X}} = 7$$

وتباين $\bar{X}$ ونرمز له $\sigma^2_{\bar{x}}$ والانحراف المعياري $\sigma_{\bar{x}}$ ويحسب كالتالي :

$$\sigma^2_{\bar{X}} = \sum_{\bar{x}} (\bar{x} - \mu_{\bar{x}})^2 f(\bar{x})$$

$$\sigma^2_{\bar{X}} = (4-7)^2\times0.1 + (5-7)^2\times0.1 + (6-7)^2\times0.2 + (7-7)^2\times0.2$$

$$+ (8-7)^2\times0.2 + (9-7)^2\times0.1 + (10-7)^2\times0.1 = 3$$

ويكون الانحراف المعياري $\sigma_{\bar{x}}$ (أو الخطأ المعياري) كالتالي :

$$\sigma_{\bar{X}} = \sqrt{3} = 1.732$$

يلاحظ مما سبق أن متوسط توزيع المعاينة للمتغيّر $\bar{X}$ يساوي متوسط المجتمع μ ولكن الانحراف المعياري لتوزيع المعاينة لـ $\bar{X}$ (الخطأ المعياري) يكون أصغر من الانحراف المعياري للمجتمع σ .

نظرية (١٠ - ١)

لعيّنات عشوائية ذات حجم n مأخوذة من مجتمع له متوسط μ وانحراف معياري

σ فإن التوزيع العيني النظري للمتوسط $\bar{X}$ يكون له توقع (متوسط) $E(\bar{X})$ مساويًا لمتوسط المجتمع μ وانحراف معياري $\sigma_{\bar{x}}$ (وأحيانًا يسمى خطأ معياريًا) يعطى كالتالي :

$$\sigma_{\bar{x}} = \begin{cases} \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}} \\ \\ \dfrac{\sigma}{\sqrt{n}} \end{cases}$$

(i) إن كان حجم المجتمع N صغير أو منتهٍ حيث N صغيرة مقارنة بـ n عندما $\frac{n}{N}$ أكبر من 5 %

(ii) السحب بدون إرجاع

(١) إن كان حجم المجتمع غير منتهٍ

(٢) إن كان N كبيرة مقارنة بـ n أي أن النسبة $\frac{n}{N}$ أقل من 5 %

(٣) السحب بإرجاع .

ملاحظــة

١ ـ المقدار $\frac{N-n}{N-1}$ يسمى بمعامل التصحيح ويقترب من الواحد الصحيح كلما كبرت N مقارنة بـ n . هذا المعامل سبق أن ظهر عند إيجاد التباين للتوزيع فوق الهندسي .

٢ ـ المعاينة بالإحلال من مجتمع منتهٍ تساوي تمامًا المعاينة بدون إحلال في مجتمع غير منتـهٍ. لأن المعاينة بالإحلال من مجتمع منتهٍ يمكن اعتباره من الناحية النظرية مجتمعًا غير منتهٍ، حيث إن أي عدد من العيّنات يمكن سحبه بدون أن يستنفد المجتمع، ويظل حجمه كما هو. مثال ذلك سحب عينة n من الكرات بالإحلال من وعاء به عدد منتهٍ N كرة من اللونين الأبيض والأسود. فإن نسبة كل من اللونين تظل ثابتة في الوعاء وهذا يحدث أيضًا عندما يكون السحب بدون إحلال ولكن في مجتمع غير منتهٍ.

مثال (١٠ ـ ٣)

أحد المجتمعات يتكون من العناصر التالية {1,2,3,4,5,6} (مثلاً رمي حجر نرد). أوجد التوزيع العيني لـ $\bar{x}$ للعيّنات ذات الأحجام التالية 4 و 3 و 2 بدون إرجاع ثم تحقق في كل حالة من الحالات السابقة من العلاقتين:

$$\mu_{\bar{x}} = \mu \ , \quad \sigma^2_{\bar{x}} = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

الحـــل

بحساب متوسط المجتمع μ وانحرافه المعياري σ وجد الناتج كالتالي:

$$\mu = 3.5 \quad , \quad \sigma = 1.708$$

أولاً: العينات ذات الحجم $n = 2$ يوجد هناك عدد $\binom{6}{2}$ يساوي 15 عينة بيانها يوضح بالجدول التالي:

متوسط العيّنة	العيّنات الممكنة	متوسط العيّنة	العيّنات الممكنة
4	{2,6}	1.5	{1,2}
3.5	{3,4}	2	{1,3}
4	{3,5}	2.5	{1,4}
4.5	{3,6}	3	{1,5}
4.5	{4,5}	3.5	{1,6}
5	{4,6}	2.5	{2,3}
5.5	{5,6}	3	{2,4}
		3.5	{2,5}

ومن النتائج السابقة يمكن إيجاد توزيع $\bar{X}$ كما هو موضح بالجدول التالي:

$\bar{X}$	$f(\bar{X}) = P(\bar{X} = \bar{n})$
1.5	1/15
2	1/15
2.5	2/15
3	2/15
3.5	3/15
4	2/15
4.5	2/15
5	1/15
5.5	1/15

من هذا التوزيع السابق يمكننا حساب ($\mu_{\bar{X}}$) كالتالي:

$$\mu_{\bar{X}} = \sum_{i=1}^{9} \bar{x}_i f(\bar{x}_i)$$

$$= 1.5 \times \frac{1}{15} + 2 \times \frac{1}{15} + 2.5 \times \frac{2}{15} + 3 \times \frac{2}{15} + 3.5 \times \frac{3}{15}$$

$$+ 4 \times \frac{2}{15} + 4.5 \times \frac{2}{15} + 5 \times \frac{1}{15} + 5.5 \times \frac{1}{15}$$

$$= 3.5$$

وهو متوسط المجتمع μ نفسه أي أن:

$$\mu = \mu_{\bar{X}} = 3.5$$

ولحساب الخطأ المعياري $\sigma_{\bar{X}}$ نحسب أولاً $\sigma^2_{\bar{X}}$ كالتالي:

$$\sigma^2_{\bar{X}} = \sum_{i=1}^{9} (\bar{x}_i - 3.5)^2 f(\bar{x}_i)$$

$$= (1.5 - 3.5)^2 \frac{1}{15} + (2 - 3.5)^2 \frac{1}{15} + (2.5 - 3.5)^2 \frac{2}{15}$$

$$+ (3-3.5)^2 \frac{2}{15} + (3.5-3.5)^2 \frac{3}{15} + (4-3.5)^2 \frac{2}{15}$$

$$+ (4.5-3.5)^2 \frac{2}{15} + (5-3.5)^2 \frac{1}{15} + (5.5-3.5)^2 \frac{1}{15}$$

$$= 1.16666$$

وتكون $\sigma_{\bar{x}}$ كالتالي:

$$\sigma_{\bar{x}} = \sqrt{1.16666} = 1.0801$$

ثانيًا: العيّنات ذات الحجم $n=3$ فإنه يوجد عدد $\binom{6}{3}$ يساوي 20 عيّنة مختلفة بيانها كالتالي:

متوسط العيّنة $\bar{x}$	العيّنات الممكنة	متوسط العيّنة $\bar{x}$	العيّنات الممكنة
3	{2,3,4}	2	{1,2,3}
3.33	{2,3,5}	2.33	{1,2,4}
3.66	{2,3,6}	2.66	{1,2,5}
3.66	{2,4,5}	3	{1,2,6}
4	{2,4,6}	2.66	{1,3,4}
4.33	{2,5,6}	3	{1,3,5}
4	{3,4,5}	3.33	{1,3,6}
4.33	{3,4,6}	3.33	{1,4,5}
4.66	{3,5,6}	3.66	{1,4,6}
5	{4,5,6}	4	{1,5,6}

ومن الجدول السابق يمكن صياغة توزيع $\bar{x}$ التكراري النسبي في الجدول التالي:

$\bar{x}$	$f(\bar{x})$	$\bar{x}$	$f(\bar{x})$
2	1/20	3.66	3/20
2.33	1/20	4	3/20
2.66	2/20	4.33	2/20
3	3/20	4.66	1/20
3.33	3/20	5	1/20

ومن هذا التوزيع بالجدول السابق نوجد $\mu_{\bar{X}}$ والتباين $\sigma_{\bar{X}}$ كالتالي:

$$\mu_{\bar{X}} = \sum_{i=1}^{10} \bar{x}_i \, f(\bar{x}_i)$$

$$= 2 \times \frac{1}{20} + 2.33 \times \frac{1}{20} + 2.66 \times \frac{2}{20} + \ldots + 5 \times \frac{1}{20}$$

$$= 3.4999997$$

$$\sigma^2_{\bar{X}} = (2 - 2.5)^2 \frac{1}{20} + \ldots + (5 - 3.5)^2 \frac{1}{20} = \frac{11.666665}{20}$$

$$= 0.5833332$$

ويكون الانحراف المعياري $\sigma_{\bar{X}}$ كالتالي:

$$\sigma_{\bar{X}} = \sqrt{0.5833332}$$

$$= 0.76376$$

ثالثًا : العيّنات ذات الحجم n = 4 فإنه يوجد عدد $\binom{6}{4}$ يساوي 15 عينة كالتالي :

متوسط العيّنات $\bar{x}$	العيّنات الممكنة	متوسط العيّنات $\bar{x}$	العيّنات الممكنة
3.75	{1,3,5,6}	2.5	{1,2,3,4}
4	{1,4,5,6}	2.75	{1,2,3,5}
3.5	{2,3,4,5}	3	{1,2,3,6}
3.75	{2,3,4,6}	3	{1,2,4,5}
4	{2,3,5,6}	3.25	{1,2,4,6}
4.25	{2,4,5,6}	3.5	{1,2,5,6}
4.5	{3,4,5,6}	3.25	{1,3,4,5}
		3.5	{1,3,4,6}

ومما سبق يمكن تكوين جدول توزيع المعاينة لـ $\bar{X}$ كالتالي :

$f(\bar{x})$	$\bar{x}$	$f(\bar{x})$	$\bar{x}$
2/15	3.75	1/15	2.5
2/15	4	1/15	2.75
1/15	4.25	2/15	3
1/15	4.50	2/15	3.25
		3/15	3.50

ومن هذا التوزيع نوجد $\mu_{\bar{X}}$ كالتالي :

$$\mu_{\bar{X}} = \sum_{i=1}^{9} \bar{x}_i f(\bar{x}_i)$$

$$= 2.5 \times \frac{1}{15} + \ldots + 4.5 \times \frac{1}{15}$$

$$= \frac{52.5}{15} = 3.5$$

$$\sigma^2_{\bar{x}} = (2.5-3.5)^2\frac{1}{15} + \ldots + (4.5-3.5)^2\frac{1}{15}$$

$$= \frac{4.375}{15}$$

$$= 0.29166$$

$$\sigma_{\bar{x}} = \sqrt{0.29166}$$

$$= 0.54006$$

ويمكن إيجاز ما سبق في الجدول التالي :

حجم العينة n	$\mu_{\bar{x}}$	$\sigma_{\bar{x}}$
2	3.5	1.0802
3	3.5	0.7638
4	3.5	0.5401

وباستخدام النظرية السابقة نجد أن :

عندما n = 2 يكون :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$= \frac{1.708}{\sqrt{2}}\sqrt{\frac{4}{5}} = 1.0802$$

عندما n = 3 يكون :

$$\sigma_{\bar{x}} = \frac{1.708}{\sqrt{3}}\sqrt{\frac{3}{5}} = 0.7638$$

وعندما n = 4 يكون :

$$\sigma_{\bar{x}} = \frac{1.708}{\sqrt{4}}\sqrt{\frac{2}{5}} = 0.5401$$

وهي النتائج نفسها التي حصلنا عليها في الجدول السابق.

تمـرين

أعد حل المثال (١٠ ـ ٣) عندما تكون المعاينة بالإحلال. ثم تحقق من العلاقتين:

$$\mu_{\bar{x}} = \mu \quad , \quad \sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

(١٠ ـ ٥) نظـرية النهايـة المركزيـة

Central limit theorem

وتعتبر هذه النظرية أساسية في علم الإحصاء والتي تعطى التوزيع العيني للمتوسط وهي كالتالي:

إذا فرض أننا أخذنا عيّنة حجمها n من جتمع له متوسط μ وانحراف معياري σ .

فإن المتغيّر العشوائي $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ له الخصائص التالية:

(١) يكون له (تقريبًا) التوزيع الطبيعي القياسي إذا كانت $n \geqq 30$ مهما كان توزيع المجتمع.

(٢) يكون له (تمامًا) التوزيع الطبيعي القياسي إذا كان توزيع المجتمع طبيعيًا مهما كان حجم العيّنة.

ملاحظـة

ا ـ إذا كانت $n \geqslant 30$ فإن $\bar{X} \sim N(\mu\,,\ \frac{\sigma^2}{n})$

ب ـ إذا كان المجتمع طبيعيًا فإن $\bar{X} \sim N(\mu\,,\ \frac{\sigma^2}{n})$ مهما كانت n .

مثال (١٠ ـ ٣)

فتائل حريرية متوسط نقطة قطعها μ يساوي 25 رطلاً بانحراف معياري σ يساوي 0.5 رطلاً. اختيرت عيّنة عشوائية حجمها 50 فتيلة وذلك لإيجاد نقطة قطعها. ما هو احتمال أن متوسط نقطة القطع للعيّنة $\bar{X}$ سيكون ما بين (25.1 و 24.9) رطلاً؟

الحـــل

المطلوب إيجاد احتمال المقدار:

$$P(24.9 \leqslant \bar{X} \leqslant 25.1)$$

حجم العيّنـة $n = 50$ وبـاستخدام نظرية النهاية المركزية فإن $\bar{X}$ يتوزع تقريبًا توزيعًا طبيعيًا بمعالم μ ، $\frac{\sigma}{\sqrt{n}}$ ويكتب عادة كالتالي:

$$\bar{X} \sim N(25, 0.005)$$

وعليـه فـإن

$$P(24.9 \leqslant \bar{X} \leqslant 25.1) =$$

$$= P\left(\frac{24.9 - 25}{0.07} \leqslant Z \leqslant \frac{25.1 - 25}{0.07}\right)$$

$$= P(-1.43 \leqslant Z \leqslant 1.43)$$

$$= \Phi(1.43) - \Phi(-1.43)$$

$$= 0.8473$$

وهذا يفسّر كالتالي

لنفرض أننا فحصنا عيّنة عشوائية مكوّنة من 50 فتيلة. عندئذ يمكن القول إنه باحتمال حوالي 85% أن متوسط نقطة القطع للعينة $\bar{X}$ سيكون اختلافه حوالي 0.1 رطلاً من المتوسط الحقيقي 25 رطلاً.

مثال (١٠ ـ ٤)

نسبة الذكاء في مجتمع ما يكون لها توزيع طبيعي بمتوسط 100 وانحراف معياري

10 . فإذا اخترنا عيّنة عشوائية حجمها 16 فردًا من هذا المجتمع فما هو احتمال أن يقع متوسط العيّنة ($\bar{X}$) بين 95 و 105 .

الحـــل

المطلوب إيجاد احتمال المقدار: $P(95 \leqslant \bar{X} \leqslant 105)$

$$P(95 \leqslant \bar{X} \leqslant 105) = P\left(\frac{95-100}{10/\sqrt{16}} \leqslant Z \leqslant \frac{105-100}{10/\sqrt{16}}\right)$$

$$= P(-2 \leqslant Z \leqslant 2)$$

$$= \Phi(2) - \Phi(-2)$$

$$= 0.97725 - 0.02275$$

$$= 0.95450$$

أي إذا أخذنا عيّنة عشوائية حجمها 16 فرد فقط فإنه لدينا أكثر من 95% احتمالاً أن تكون $\bar{X}$ مختلفة عن 100 بحوالي 5 وحدات فقط.

(١٠ - ٦) تمـــارين

(١) i) كم عيّنـة عشوائية مختلفة حجمها n=2 يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكونة من:

$$N = 25 \quad , \quad N = 15 \quad , \quad N = 10 \quad , \quad N = 6$$

ii) كم عيّنـة عشوائية مختلفة حجمها n=3 يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكوّنة من N = 20 , N = 25 , N = 50 ثم أوجد احتمال اختيار كل عيّنة.

iii) ما هو احتمال اختيار العيّنات الممكنة التالية (بدون إرجاع)

ا) اختيار عيّنة عشوائية حجمها n=4 من مجتمع يتكون من N=12 .

ب) عيّنة عشوائية حجمها n=5 من مجتمع يتكون من N=22 .

(٢) i) مجتمع منتهٍ يتضمن N عنصرًا سحبنا منه عيّنة بإرجاع ثم بدون إرجاع حجمها n متى نقول إن هذه العيّنة عشوائية .

ii) لدينا المجتمع المنتهِ 12 و 10 و 8 و 6 . أكتب له كل العيّنات العشوائية الممكنة بإرجاع ثم بدون إرجاع المؤلفة من عنصرين من عناصر هذا المجتمع . ثم أوجد توزيع $\bar{X}$ متوسط العيّنة واحسب متوسط وتباين هذا التوزيع .

(٣) سحبنا عيّنة عشوائية حجمها $n=64$ من مجتمع كبير جدًا متوسطه $\mu=12$ وانحرافه المعياري $\sigma=4$ وليكن $\bar{X}$ هو متوسط هذه العيّنة .

i) ما هو توقع $\bar{X}$ وتباين $\bar{X}$.

ii) أحسب احتمال أن يختلف متوسط العيّنة عن μ بأقل من 0.8 .

(٤) i) كم عدد العيّنات (بدون إرجاع) والتي حجم كل منها $n=3$ يمكن تكوينها من المجتمع الذي يتكون من العناصر a, b, c, d, e . ثم أوجد احتمال اختيار أي من العيّنات الممكنة .

ii) ما هو احتمال اختيار أي عنصر معين b مثلاً في العيّنة المسحوبة .

iii) ما هو احتمال اختيار أي عنصرين معينين c و d مثلاً .

(٥) إذا سحبنا بدون إرجاع عيّنة عشوائية حجمها 2 من مجتمع محدود يحتوي على الأعداد التالية : 5, 6, 7, 8, 9, 10 .

i) أثبت أن متوسط المجتمع هو 7.5 وانحرافه المعياري هو 0.35112 .

ii) أوجد كل العيّنات الممكنة لهذا المجتمع وأوجد متوسطها .

iii) مستخدمًا النتائج السابقة في (ii) ووضع لكل عيّنة احتمال $\frac{1}{15}$ أنشىء توزيع المعاينة للمتوسط وذلك للعيّنات ذات الحجم 2 من هذا النوع .

iv) أحسب المتوسط والانحراف المعياري للتوزيع الذي أنشىء في (iii) .

(٦) سحبنا عيّنات عشوائية حجمها 3 من المجتمع في التمرين (5) السابق حل الفقرات (ii) ، (iii) ، (iv) .

(٧) ما هي قيمة معامل التصحيح للمجتمع المحدود عندما تكون:

a) $n = 5$, $N = 200$
b) $n = 10$, $N = 300$
c) $n = 15$, $N = 45$

(٨) متوسط عيّنة عشوائية حجمها $n = 36$ يستخدم لتقدير متوسط مجتمع غير محدود والذي له انحراف معياري $\sigma=9$. ما هو احتمال أن يكون أقل من 4.5 إذا استخدمنا:

ا) نظرية تشيبشف.
ب) نظرية النهاية المركزية.

(٩) متوسط عيّنة عشوائية حجمها $n = 25$ يستخدم كتقدير المتوسط مجتمع كبير له انحراف معياري $\sigma=2.4$. ما هو احتمال أن يكون الخطأ أقل من 1.2 إذا استخدمنا:

ا) نظرية تشيبشف.
ب) نظرية النهاية المركزية.

(١٠) عند أخذ العيّنات من مجتمع غير محدود. ماذا يطرأ على تغيير الخطأ المعياري للمتوسط ($\sigma_{\bar{x}}$) بمعنى حجم الخطأ الذي نتعرض إليه عندما نستعمل $\bar{X}$ كتقدير لـ μ .

i) إذا زيد حجم العيّنة من 60 إلى 240 ؟
ii) إذا زيد حجم العيّنة من 200 إلى 450 ؟
iii) إذا زيد حجم العيّنة من 25 إلى 225 ؟
iv) إذا نقص حجم العيّنة من 60 إلى 40 ؟

(١١) ما هي قيمة معامل التصحيح للمجتمع المحدود عندما تكون:

i) $n = 5$, $N = 200$

ii) $n = 10$, $N = 300$

iii) $n = 100$, $N = 5000$

(١٢) إذا علمت أن X موزعة توزيعًا طبيعيًا، وسطه 25 ، وانحرافه المعياري 8 . أحسب احتمال أن يكون وسط العينة $\bar{X}$. المبني على عينة حجمها 16 .

ا) أقل من 26 ب) أكبر من 31

جـ) أقل من 21 د) بين 29 و 28 .

(١٣) أرسم، على الورقة نفسها شكلا لمنحنى طبيعي، وسطه عشرة، وانحرافه المعياري 2 ، وشكلا آخر لمنحنى توزيع المتوسط المناظر للعيّنات التي حجمها 9 .

(١٤) ماذا يكون عليه شكل منحنى $\bar{X}$ في مسألة (١٣) لو كان حجم العيّنة يساوي 36 .

(١٥) إذا علمت أن X موزعة توزيعًا طبيعيًا وسطه 20 وانحرافه المعياري 4 أحسب احتمال أن يكون وسط العيّنة $\bar{X}$ التي حجمها 4 في الحالات التالية :

i) أكبر من 21 ii) أكبر من 19.5

iii) بين 21 و 19 iv) أكبر من 22 .

(١٦) إذا كان الانحراف المعياري لأوزان أطفال المرحلة الأولى هو 7 أرطال. فما هو احتمال أن يختلف الوزن لمتوسط عيّنة عشوائية من 100 من أمثال هؤلاء الأطفال بأكثر من 1 رطل عن الوزن المتوسط لجميع الأطفال.

(١٧) نظام غذائي للتسمين يراد تطبيقه على عيّنة عشوائية من 25 كتكوتًا مأخوذة من حظيرة. فإذا كان من المتوقع أن يكون الانحراف المعياري لزيادة الوزن خلال فترة شهر حوالي $\frac{1}{2}$ أوقية فما احتمال أن يختلف وسط هذه العيّنة بأكثر من $\frac{1}{2}$ أوقية

عن وسط كتـاكيت الحظيرة كلها إذا سارت على هذا النظام الغذائي الجديد الذي لا يفضّل عن النظام القديم؟

(١٨) اذكر ثلاثة أسباب لاختيار العيّنة.

(١٩) اقترح طريقة لأخذ عيّنة من 100 طالب من بين طلبة الجامعة.

١١

تقدير معالم المجتمع واختبارات الفروض

Estimation and Testing Hypotheses

● مقدمة ● التقدير بنقطة ● التقدير بفترة ● القيمة العظمى للخطأ في التقدير ● حجم العيّنة ● تقدير فترة الثقة للمتوسط ● تقدير فترة الثقة للنسبة R ● اختبارات الفروض الإحصائية ● الاختبارات المعنوية ● الخطأ من النوع الأول (α) والخطأ من النوع الثاني (β) ● اختبار الفرضيّات للمتوسط μ ● اختبار الفرضيات للنسبة R ● تمارين

(١١ ـ ١) مقدمة

سبق لنا تعريف المجتمع الإحصائي والعيّنة الإحصائية العشوائية، والأسباب التي تؤدي إلى أخذ العيّنات لدراسة المجتمع بدلاً من الحصر الشامل. وكذلك درسنا بعض التوزيعات الاحتمالية مثل التوزيع ذي الحدين والتوزيع فوق الهندسي وتوزيع بواسون، والتوزيع الطبيعي. وهي نماذج احتمالية تمثل توزيع أي ظاهرة في المجتمع مثل ظاهرة وصول عدد من المكالمات الهاتفية إلى المجمّع الهاتفي، أو ظاهرة الطول، أو الوزن أو الذكاء للإنسان. . . الخ. والتوزيعات السابقة تحتوي على معالم مثل n.p

للتـوزيع ذي الحـدين، λ لتـوزيع بواسون، σ ، μ للتوزيع الطبيعي ويتحدد شكل التوزيع تمامًا بتحديد قيم هذه المعالم.

وعادة ما تكون هذه المعالم مجهولة، ويرغب الباحثون في تقدير مثل هذه المعالم من البيانات المأخوذة من العيّنات العشوائية. والمطلوب الآن البحث عن أنسب دالة أو إحصاءة (statistic) تعتمد على قيم العيّنة، ولا تعتمد على المعالم المجهولة للمجتمع كتقدير هذه المعالم. وهناك طريقتان أساسيتان لتقدير معالم المجتمع المجهولة هما التقدير بنقـطة والتقدير بفترة ونوضح كل منهما فيما يلي ثم نتناول بعد ذلك الفرع الثاني من الاستدلال الإحصائي وهو اختبارات الفروض.

(١١ - ٢) التقدير بنقطـة

يحتـوي المجتمع الإحصائي عادة على معالم تكون غير معلومة كمتوسطه μ أو انحرافه σ أو نسبة معينة R . . . الخ ويمكن إيجاد تقديرات لهذه المعالم من البيانات المـأخـوذة من عيّنـة عشوائية وذلك بحساب ما يسمى بالإحصاءات (وهي دوال في المشـاهـدات) فمثـلاً متوسط العيّنة العشوائية $\bar{X}$ يستخدم كتقدير لمتوسط المجتمع μ وكـذلـك الانحـراف المعياري S يستخـدم كمقـدر للانحراف المعياري للمجتمع σ وهكذا. . . وتسمى هذه التقديرات التقدير بنقطة لأنها قيمة وحيدة محسوبة من العيّنة.

ومثال لذلك إذا أردنا إيجاد قيمة تقديرية لمتوسط طول الطلّاب الدارسين لمقرر 101 إحص بجامعة الملك سعود في أحد الفصول الدراسية وأخذنا عيّنة عشوائية من هؤلاء الطلّاب ووجدنا أن متوسط طول هذه العيّنة مثلاً هو 160 cm $(\bar{x} = 160)$ فإنه يمكن القـول إن القيمـة التقديرية للمتوسط μ لأطوال الطلّاب الدارسين لمقرر 101 إحص هو 160 cm . وكذلك إذا أخذنا عيّنة عشوائية من إنتاج إحدى المصانع ووجدنا أن نسبة المعيب في هذه العيّنة هو % 5 فإنه يمكن القول إن القيمة التقديرية لنسبة المعيب في المصنع هي % 5 والتقدير بنقطة له خصائص يجب أن يحققها حتى يكون تقدير جيد، ولن نتعرض لهذه الخصائص لأنها خارج نطاق هذا الكتاب.

(١١ ـ ٣) التقـدير بفـترة

التقدير بفترة لإحدى معالم المجتمع المجهولة .... $\mu, \sigma, R,$ هي عبارة عن إيجاد فترة تحدد بقيمتين تحسب من مشاهدات العيّنة العشوائية المأخوذة من هذا المجتمع محل الدراسة، ونتوقع احتواء هذه الفترة على معلمة المجتمع باحتمال معين $(1 - \alpha)$ حيث عادة α تأخذ قيمًا صغيرة مثل ... ,0.1, 0.05, 0.01 ويمكن أيضًا إيجاد دقة التقدير بفترة للمعلمة. وكلما كان طول الفترة صغيراً زادت دقة التقدير. لذلك سميت بتقدير فترة الثقـة. فإذا كان مثـلاً درجـة الدقة في الخطأ بين متوسط المجتمع μ ومتوسط العيّنة العشوائية $\bar{x}$ هو المقدار الموجب ε فإنه يمكن تحديد حد أدنى للاحتمال يكتب كالتالي:

$$P(|\mu - \bar{x}| \leqslant \varepsilon) \geqslant 1 - \alpha \quad \text{............} \quad (١١ - ١)$$

ويمكن كتابة العلاقة (١١ ـ ١) كالتالي:

$$P(\bar{x} - \varepsilon \leqslant \mu \leqslant \bar{x} + \varepsilon) \geqslant 1 - \alpha \quad \text{............} \quad (١١ - ٢)$$

والعلاقة (١١ ـ ٢) تمثل فترة ثقة $(\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ للمعلمة المجهولة μ باحتمال لا يقل عن $(1 - \alpha)$. والمقدار $(1 - \alpha)$ يسمى درجة الثقة فإذا كانت قيم α هي 0.1, 0.05, 0.01 فإن درجات الثقة المناظرة هي على الترتيب % 99 ,% 95 ,% 90 .

وسوف نقوم باستعراض كيفية تقدير فترة الثقة التالية لبعض المعالم المجهولة للمجتمع.

١ ـ تقدير فترة الثقة للمتوسط μ لمجتمع تباينه σ^2 معروف أو غير معروف للعيّنات الكبيرة.

٢ ـ تقدير فترة الثقة للمتوسط μ لمجتمع ذات توزيع طبيعي تباينه σ^2 معروف أو غير معروف للعيّنات الصغيرة.

٣ ـ تقدير فترة الثقة للنسبة R .

٤ ـ تقدير فترة الثقة للفرق بين متوسطين $(\mu_1 - \mu_2)$ لكل من العيّنات الكبيرة والعيّنات الصغيرة.

٥ ـ تقدير فترة الثقة للفرق بين نسبتين $(R_1 - R_2)$.

وقبل التعرض لدراسة كيفية حساب تقدير فترات الثقة للحالات السابقة. سوف نستعرض فيما يلي ما يسمى بالقيمة العظمى في خطأ التقدير، والتي تساعدنا في إيجاد حجم العيّنة عند مستوى دقة معين (α)، وكذلك في إيجاد فترات الثقة السابق ذكرها.

(١١ ـ ٤) القيمة العظمى للخطأ في التقدير

لقد سبق دراسة توزيع المعاينة لمتوسط العينة $\bar{X}$ ووجد أن الخطأ المعياري له ($\sigma_{\bar{X}}$) يساوي $\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ وذلك عندما يكون حجم العينة n مأخوذًا من مجتمع صغير أو محدود حجمه N. أو يساوي $\frac{\sigma}{\sqrt{n}}$ في المجتمعات الكبيرة جدًا أي عندما يكون حجم العيّنة n يمثل نسبة صغيرة جدًّا من المجتمع الذي حجمه N. ولقد سبق أيضًا دراسة نظرية النهاية المركزية. والتي تقول إن توزيع المعاينة للمتوسط يقترب من التوزيع الطبيعي وأنه يمكن أن يؤكد باحتمال قدره $(1 - \alpha)$ بأن متوسط العينة $\bar{X}$ يختلف عن متوسط المجتمع μ بمقدار يقل عن $z_{\alpha/2}$ من الخطأ المعياري $\sigma_{\bar{X}}$. ويمكن التعبير عما سبق بالعبارة الاحتمالية التالية:

$$P(|\bar{x}-\mu| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha \qquad (١١ - ٣)$$

وذلك في حالة المجتمعات ذات الحجم الكبير أو اللانهائية. والمقدار $|\bar{X}-\mu|$ يسمى خطأ التقدير (ونرمز له بالرمز E) وذلك عندما يزيد مقدار الخطأ في التقدير إلى نهايته العظمى. فمن المعادلة (١١ ـ ٣) نجد أن خطأ التقدير يأخذ القيمة التالية:

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad (١١ - ٤)$$

أحيانًا تسمى القيمة العظمى في خطأ التقدير بالدقة في التقدير ونوضح ذلك بالمثال التالي.

مثال (١١ ـ ١)

في دراسة عن تلوّث الهواء بأوكسيد الكبريت المنبعث من أحد المصانع . سحبت عيّنة مكوّنة من قراءات 80 يومًا، وحُسب متوسط العيّنة فوجد أنه يساوي 18.85 طنًّا بانحراف معياري قدره 5.55 طنًّا.

أحسب وباحتمال (0.95) مقدار القيمة العظمى للخطأ في التقدير E .

الحـــل

من المعادلة (١١ ـ ٤) السابقة نجد أن :

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\therefore\ E = 1.96 \times \frac{5.55}{\sqrt{80}} = 1.22 \quad \text{طنًّا}$$

(١١ ـ ٥) حجـم العيّنـة

عنـد تحديد مقـدار الـدقـة المـطلوبـة (أو القيمة العظمى للخطأ في التقدير) عنداحتمال معين $(1-\alpha)$ يمكن تحديد حجم العيّنة n من المعادلة (١١ ـ ٤) ويكون كالتالي :

$$n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2 \qquad \text{(١١ ـ ٥)}$$

مثال (١١ ـ ٢)

يرغب صاحب مصنـع في تقدير حجم العيّنة n . حتى يمكنه التأكد تأكدًا معقولاً من أن تقديره وباحتمال 0.95 مثلاً لن يكون مخطئًا بأكثر من 5 وحدات معيبة، إذا علم أن الانحراف المعياري يساوي 20 وحدة.

الحـــل

حجم العيّنة المطلوب يحسب بالمعادلة التالية :

$$n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2$$

حيث

$$E = 5 \quad ، \quad \sigma = 20 \quad ؛ \quad z_{\alpha/2} = 1.96$$

$$n = \left(\frac{1.96 \times 20}{5}\right)^2 \simeq 62$$ وحدة

والصيغة السابقة (١١ ـ ٥) تمكننا من تحديد مدى كبر العيّنة التي نحتاج إليها لكي يقدر متوسط المجتمع μ لأي درجة نرغبها شريطة أن يكون الانحراف المعياري للمجتمع معلومًا.

(١١ ـ ٦) تقدير فترة الثقة للمتوسط

(١١ ـ ٦ ـ ١) في العيّنات الكبيرة

في المعادلة (١١ ـ ٣) يمكننا أن نقول إنه باحتمال $(1 - \alpha)$ أن:

$$|\bar{x} - \mu| \leq z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

أي أن:

$$-z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

أي أن

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \qquad (١١ ـ ٦)$$

وهذا يعني بأن متوسط المجتمع μ واقع داخل الفترة الممتدة من الحد الأعلى $\bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ إلى الحد الأدنى $\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ ويسمى هذا بتقدير فترة الثقة عند مستوى معنوي أو بدرجة ثقة قدرها $100(1 - \alpha)$. فإذا كانت $(\alpha = 0.05)$ فإن درجة الثقة هي 95% . وإذا كانت $(\alpha = 0.01)$ فإن درجة الثقة تكون (99%) وهكذا.

مثال (١١ ـ ٣)

من مثال (١١ ـ ١) أوجد بدرجة ثقة (95%) تقدير فترة الثقة للمتوسط الحقيقي لانبعاث أكسيد الكبريت في الجو من أحد المصانع.

الحـــل

تقدير فترة الثقة المطلوبة تعطى من المعادلة (١١ ـ ٦) كالتالي:

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + _{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$18.85 - 1.96 \frac{5.55}{\sqrt{80}} < \mu < 18.85 + 1.96 \frac{5.55}{\sqrt{80}}$$

$$17.63 < \mu < 20.07$$

أي أن:

تقدير الحد الأعلى لإنبعاث أكسيد الكبريت في الجو = 20.07 طنًا.

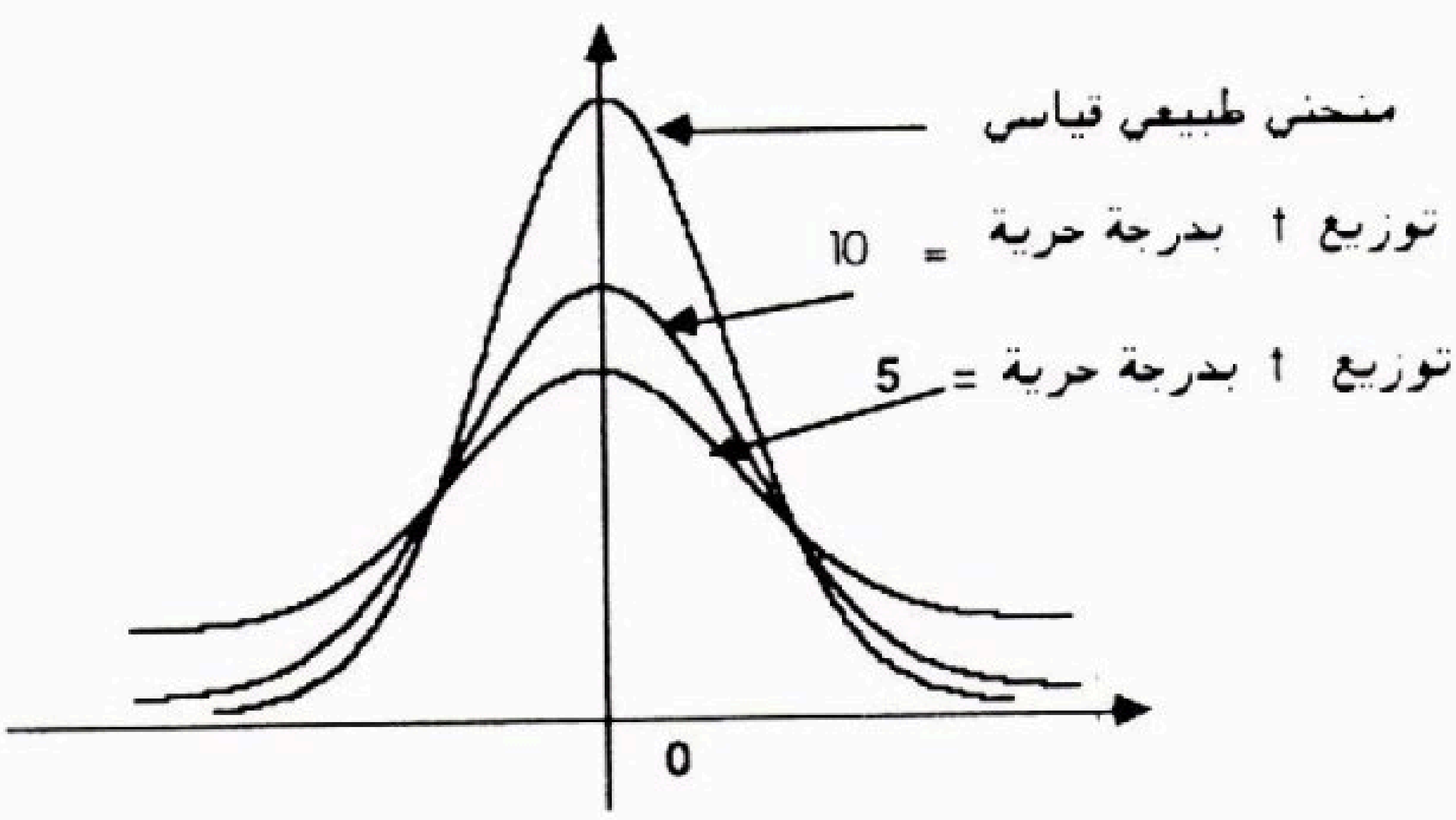
تقدير الحد الأدنى لإنبعاث أكسيد الكبريت في الجو = 17.63 طنًا.

(١١ ـ ٦ ـ ٢) في العيّنات الصغيرة

لقد سبق أن ذكرنا أن توزيع المعاينة للمتوسط يقترب من التوزيع الطبيعي وذلك عندما يكون الانحراف المعياري (σ) للمجتمع معلومًا أو غير معلوم للعيّنات الكبيرة أو معلومًا للعيّنات الصغيرة من مجتمع ذا توزيع طبيعي ولكن عندما يكون الانحراف المعياري للمجتمع الطبيعي (σ) مجهولاً ونستعيض عنه بالانحراف المعياري المقدّر من العيّنة (s) فإن توزيع المعاينة توزيع آخر متماثل ولكن يختلف عن التوزيع الطبيعي. ولقد اكتشف هذا التوزيع العالم الألماني جوست (W.S.Gosset) في أوائل القرن العشرين ونشره تحت اسم مستعار (توزيع t) أو توزيع طالب. وهذا التوزيع يشبه التوزيع الطبيعي ولكن قمته أكثر انخفاضًا عن قمة التوزيع الطبيعي ويتقارب طرفيه من الصفر عند $\infty +$, $\infty -$ أبطأ من تقارب طرفيْ التوزيع الطبيعي ويقترب من التوزيع الطبيعي القياسي كلما كبر حجم العيّنة n أو زادت درجات الحرية v حيث ($v = n - 1$) وينطبق على التوزيع الطبيعي عندما يصبح $v = 30$. ويمكن كتابة دالة كثافة توزيع t كالتالي:

$$f(t) = K\left(1 + \frac{t^2}{v}\right)^{\frac{1-v}{2}} \quad ; \quad -\infty < t < \infty .$$

وحيث K ثابت يعتمد على درجات الحرية v ، ليجعل المساحة تحت المنحنى تساوي الواحد الصحيح . ونوضح شكل توزيع t عند درجات حرية 5 ,10 ومنحنى التوزيع الطبيعي القياسي شكل (١١ - ١) التالي :



شكل (١١ - ١) يمثل توزيع t مقارنة بالتوزيع الطبيعي

ويمكن كتابة فترة الثقة للعيّنات الصغيرة باحتمال قدره $(1 - \alpha)$ لمتوسط المجتمع μ مثل ما تم بالنسبة للعيّنات الكبيرة في الصيغة (١١ - ٦) وتكون كالتالي :

$$\bar{x} - t_{(\alpha/2, v)} \left(\frac{s}{\sqrt{n}} \right) < \mu < \bar{x} + t_{(\alpha/2, v)} \left(\frac{s}{\sqrt{n}} \right) \qquad \text{(١١ - ٧)}$$

حيث استبدلنا $Z_{\alpha/2}$ بالقيمة $t_{\alpha/2}$. واستبدلنا الانحراف المعياري σ للمجتمع بالانحراف المعياري للعيّنة S . و $t_{\alpha/2}$ تعطى من جداول t رقم (3) بمؤخرة الكتاب حيث تعطى قيم t لكل درجة من درجات الحرية $v = n - 1$ وذلك لبعض قيم α مثل ... , 0.01 , 0.05 , 0.1 ونوضح طريقة حساب فترة الثقة للعينات الصغيرة بالمثال التالي .

مثال (١١ ـ ٤)

لدراسة اختفاء العلامات الأرضية البيضاء في وسط الطريق نتيجة المرور الكثيف. أخذت عيّنة من 8 مناطق مختلفة، ولوحظ اختفاء العلامات بعد مرور السيارات مقربة لأقرب مائة سيارة كالتالي:

167800, 136500, 108300, 126400,

133700, 162000, 149400, 142600

قدّر فترة الثقة للمتوسط μ من هذه البيانات.

الحـــل

$$\bar{x} = (167800 + 136500 + \ldots + 149400) / 8$$

$$= 140800 \qquad \text{سيـــارة}$$

$$S^2 = \frac{(167800 - 140800)^2 + (136500 - 140800)^2 + \ldots + (142600 - 140800)^2}{8 - 1}$$

أي أن:

$$S = 19200 \qquad \text{سيـــارة}$$

درجات الحرية v تعطى كالتالي:

$$v = 8 - 1 = 7$$

ونحسب $t_{(\alpha/2,y)}$ حيث نختار مستوى المعنوية α عبارة عن 0.05 أي عند درجة ثقة قدرها 0.95 من جدول t فتكون:

$$t_{(\alpha/2,v)} = t_{(0.025\,,\,7)} = 2.365$$

وبالتعويض في فترة الثقة التالية:

$$\bar{x} - t_{(\alpha/2,v)}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + t_{(\alpha/2,v)}\left(\frac{s}{\sqrt{n}}\right)$$

$$140800 - 2.365\ \frac{19200}{\sqrt{8}} < \mu < 140800 + 2.365\,\frac{19200}{\sqrt{8}}$$

أي أنه بدرجة ثقة 95% يكون تقدير فترة الثقة للمتوسط μ هي :

$$124745.85 < \mu < 156854.15$$

ومنها يكون الحد الأعلى لتقدير فترة الثقة لمتوسط عدد السيارات المارة هو 156854.15 سيارة. والحد الأدنى لتقدير فترة الثقة لمتوسط عدد السيارات المارة هو 124745.85 سيارة.

ويمكن إيجاد أقصى خطأ في التقدير E مثل ما تم في المعادلة (١١ ـ ٤) كالتالي :

$$E = t_{\alpha/2} \left(\frac{s}{\sqrt{n}} \right) \qquad \ldots\ldots\ldots \quad (١١ ـ ٨)$$

ومن المعادلة (١١ ـ ٨) يمكن إيجاد حجم العينة n عند أقصى خطأ يطلب.

مثال (١١ ـ ٥)

في تجربة على 12 من روّاد الفضاء في مجال يحاكي مجال انعدام الوزن وجد أن متوسط زيادة ضربات القلب لهم 27.33 دقة في الدقيقة بانحراف معياري قدره 4.28 دقة في الدقيقة. أحسب أقصى خطأ في تقدير المتوسط بدرجة معنوية $\alpha = 0.01$.

الحـــل

أقصى خطأ في تقدير المتوسط E يعطى من المعادلة (١١ ـ ٨) كالتالي :

$$E = t_{(\alpha/2\ ,\ \nu)} \left(\frac{s}{\sqrt{n}} \right)$$

حيث

$$\nu = n - 1 = 12 - 1 = 11$$

$$s = 4.28$$

$$\frac{\alpha}{2} = \frac{0.01}{2} = 0.005$$

من جدول t رقم (3) بمؤخرة الكتاب نجد أن :

$$t(0.005\ ,\ 11) = 3.106$$

$$E = 3.106 \left(\frac{4.28}{12} \right) = 3.84$$

وبذلك يمكن القول عند استخدام المتوسط $\bar{X} = 27.33$ دقة في الدقيقة كتقدير للمتوسط الحقيقي لزيادة الضربات فإنه يمكن أن نؤكد باحتمال 0.99 بأن أقصى خطأ في التقدير يكون أقل من 3.84 دقة في الدقيقة.

(١١ ـ ٧) تقدير فترة الثقة للنسبة R

إذا كانت الإحصاءة (r) هي نسبة النجاح في الحصول على قطعة معيبة عند سحب عيّنة حجمها n من إنتاج إحدى الآلات. أو نسبة النجاح في الحصول على طالب يدخن عند سحب عيّنة حجمها n من طلاّب الجامعة، . . . إلخ. فإن مثل هذه المجتمعات تتبع التوزيع ذي الحدين بمعلمة R ويكون التوقع والتباين للإحصاءة r ، مثل ما سبق دراسته لتوزيع المعاينة للمتوسط $\bar{X}$ كالتالي:

$$E(r) = R$$

$$V(r) = \begin{cases} \dfrac{RQ}{n} & \text{للمجتمعات الكبيرة} \\[2ex] \dfrac{RQ}{n}\,\dfrac{N-n}{N-1} & \text{للمجتمعات الصغيرة} \end{cases}$$

$$; \qquad Q = 1 - R$$

حيث N تمثل حجم المجتمع الصغير (المحدود) الذي أخذت منه العيّنة n وأن توزيع الإحصائية التالية $(r-R)/\sigma$ يقترب من التوزيع الطبيعي القياسي $N(0,1)$ وحيث σ_r تمثل الخطأ المعياري للنسبة r . وتحسب σ_r بالجذر التربيعي للتباين $V(r)$. وكما سبق أن قدرنا فترة ثقة للمتوسط μ فإنه يمكن بالمثل تقدير فترة ثقة للنسبة R وذلك في حالة المجتمعات الكبيرة كالتالي:

$$r \pm z_{\alpha/2} \sqrt{\frac{RQ}{n}} \qquad (١١ ـ ٩)$$

وللمجتمعات الصغيرة (المحددة) كالتالي :

$$r \pm z_{\alpha/2} \sqrt{\frac{RQ}{n}\left(\frac{N-n}{N-1}\right)} \qquad \text{(١١ - ١٠)}$$

ملاحظة :

عندما تكون النسبة R للمجتمع مجهولة تستبدل بالنسبة r المحسوبة من العيّنة وكذلك Q بالقيمة q حيث $q = 1 - r$ لحساب σ_r في القوانين السابقة .

مثال (١١ - ٦)

في دراسة لاختبار صلاحية إحدى الطرق لعلاج مرض ما أخذت عينة مكونة من 400 شخص ، أبدى 136 شخصًا عدم شعورهم بالراحة أثناء العلاج . قدر فترة الثقة للنسبة R عند درجة ثقة قدرها 0.95 .

الحـــل

$$r = \frac{x}{n} \quad , \quad x = 136 \quad , \quad n = 400$$

أي أن :

$$r = \frac{136}{400} = 0.34$$

$$z_{0.025} = 1.96 \quad , \quad q = 1 - r = 1 - 0.34 = 0.66$$

وتقدير فترة الثقة للنسبة R من المعادلة (١١ - ٩) السابقة كالتالي :

$$r - z_{\alpha/2}\sqrt{\frac{rq}{n}} < R < r + z_{\alpha/2}\sqrt{\frac{rq}{n}}$$

$$0.34 - 1.96\sqrt{\frac{0.34 \times 0.66}{400}} < R < 0.34 + 1.96\sqrt{\frac{0.34 \times 0.66}{400}}$$

$$0.294 < R < 0.386$$

ونحسب القيمة العظمى للخطأ في التقدير (E) في حالة النسبة R كما قد بيناه في شرحنا بالنسبة للمتوسط μ ويعطى كالتالي:

$$E = z_{\alpha/2}\sqrt{\frac{rq}{n}} \qquad \text{(١١ - ١١)}$$

مثال (١١ - ٧)

في دراسة لعيّنة مكوّنة من 150 شخصًا في أحد المصايف لمعرفة رأيهم عند تفضيلهم هذا المصيف عن غيره أجاب عدد قدره 108 منهم بأن سبب التفضيل هو دفء هذا المصيف. أحسب مقدار القيمة العظمى للخطأ في التقدير E عند درجة ثقة 0.99 .

الحـــل

حيث إن E تعطى بالعلاقة

$$E = z_{\alpha/2}\sqrt{\frac{rq}{n}}$$

$$z_{\alpha/2} = 2.58 \quad , \quad r = \frac{108}{150} = 0.72 \quad , \quad q = 0.28$$

$$E = 2.58\sqrt{\frac{(0.72)\,(0.28)}{150}} = 0.095$$

(١١ - ٧ - ١) حجم العيّنة باستخدام النسبة R

يمكن تحديد حجم العيّنة n من المعادلة (١١ - ١١) السابقة كالتالي:

$$n = r\,q\,(z_{\alpha/2}/E)^2 \qquad \text{(١١ - ١٢)}$$

ونوضح ذلك بالمثال التالي.

مثال (١١ - ٨)

لتحديد مقدار نسبة المصابين بضغط عالٍ من بين الأشخاص البالغين كان **المطلوب هو تقدير حجم العينة n حتى يمكننا التأكد باحتمال 0.99 من أن الخطأ لا يتجاوز 0.05 وذلك إذا علمنا أن قيمة r تساوي 0.2 .**

الحـــل

$$z_{\alpha/2} = 2.58 \quad , \quad E = 0.05 \quad , \quad r = 0.2 \quad , \quad q = 0.8$$

$$n = r\,q(z_{\alpha/2} / E)^2 = 0.2(0.8)\,(\frac{2.58}{0.05})^2 = 427$$ شخص

ملاحظـة

يمكن إيجاد تقدير فترات الثقة للفرق بين متوسطين $(\mu_1 - \mu_2)$ وذلك في حالة العيّنات الكبيرة $(n_1, n_2 \geqslant 30)$ والعيّنات الصغيرة $(n_1, n_2 < 30)$ عندما يكون الانحراف المعياري لكل منهما مجهولاً في كل من الحالتين السابقتين:

(ا) العينات الكبيرة:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

(ب) العيّنات الصغيرة حيث $s_1^2 = s_2^2 = \sigma^2$:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{(v,\,\alpha/2)}\, s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

حيث $v = n_1 + n_2 - 2$.

$$s^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}$$

وكذلك يمكن إيجاد تقدير فترة الثقة بين نسبتين $(R_1 - R_2)$ كالتالي:

$$(r_1 - r_2) \pm z_{\alpha/2} \sqrt{\frac{r_1(1-r_1)}{n_1} + \frac{r_2(1-r_2)}{n_2}}$$

أنظـر التمــارين.

(١١ - ٨) اختبارات الفروض الإحصائية

Test of statistical hypotheses

نتناول الآن الفرع الثاني من الاستدلال الإحصائي وهو اختبارات الفروض. يحاول الباحث في كثير من الأحيان اتخاذ قرار بشأن خواص توزيع مجتمع ما، وذلك بناء على بيانات عيّنة عشوائية اختيرت من المجتمع نفسه. فمثلاً يريد طبيب أن يختبر فعالية دواء جديد بالنسبة لعلاج مرض معيّن. أو يريد باحث في التربية أن يختبر كون منهج معيّن أكثر فائـدة من منهـج آخـر. . الـخ. وهكـذا وللوصول إلى هذه القرارات الإحصـائية (statistical decision) نقـوم عادة بوضـع فروض عن خواص المجتمع، متوسطه، انحرافه المعياري. . الخ. ونختبر هذا الفرض بناء على عينة عشوائية نختارها من المجتمـع، وهذه الفروض - التي قد تكون صحيحة، أو غير صحيحة - تسمى بالفروض الإحصائية وتنقسم الفروض إلى قسمين:

i) فروض عن معالم المجتمع (parametric)

ii) فروض عن صورة دالة التوزيع (nonparametric)

وسوف نكتفي بدراسة الفروض عن بعض معالم المجتمع فقط في هذا الكتاب.

ونبدأ بافتراض إحصائي يسمى فرض العدم (null hypothesis) (وسمي بهذا لأن الغـرض منه هو عدم قبوله أو محوه) ويرمز له بـ H_0، ويسمى الافتراض الذي يختلف عن H_0 بفرض البديل ويرمز له بـ H_1 (alternative).

وكما سبق أن وّضحنا بأن اختبار الفروض يعتمد على بيانات العيّنة وهي تعني بالإجابة على السؤال التالي: هل بيانات العيّنة متناسقة مع فرض معيّن؟ وهل تميل إلى تأكيده أو نفيه؟

فإذا فرضنا قيمة لمعلم من معالم المجتمع (مثلاً المتوسط) وأخذنا عيّنة عشوائية من المجتمع فسيكون هناك غالبًا فرق بين قيمة المعلم المفروضة والتقدير المحسوب له من بيانات العيّنة، فهل هذا الفرق راجع إلى مجرد الصدفة والخطأ المتوقع نتيجة العيّنة؟ أم أنه فرق حقيقي معنوي (significant) ، و أن هناك أسبابًا أخرى ساعدت على كبر هذا الفرق؟ فإذا وجدنا أن الفرق معنوي فإننا نميل إلى أن الفرض H_0 غير صحيح (أو على الأقل لا نقبله بناء على البيانات المتاحة). وإذا كان الفرق غير معنوي (nonsignificant) فليس هناك ما يدعونا لرفض فرض العدم H_0 .

فمثلاً إذا قذفنا قطعة نقود 20 مرة وكانت النتيجة المشاهدة هي 16 للصورة و 4 للكتابة، فإننا قطعًا نميل إلى رفض H_0 بأن العملة متّزنة، على الرغم من أننا قد نكون مخطئين في اتخاذ هذا القرار الإحصائي. ولكن إن كانت النتيجة مثلاً 11 صورة و 9 كتابة فليس هناك ما يبرر رفض H_0 .

وكمثال آخر لفرض العدم (الفرضيّة الأولية). إذا افترض أحد الباحثين بأن أطوال الذكور البالغين في إحدى البلاد هو 165 سم واختار عينة مكونة من 64 ذكرًا بالغًا. وحسب متوسط الطول لهم فكان 160 سم، وكان الانحراف المعياري لهذا المجتمع معروفًا ويساوي 5 سم. فيكون فرض العدم في هذه الحالة $H_0 : \mu = 165$ cm (أو الفرضيّة الأولية H_0). حيث نفترض عدم وجود فروق حقيقية بين متوسط المجتمع والقيمة المفروضة. والفروق المشاهدة إنما تعزى للصدفة ويقابل الفرضيّة الأولية فرضيّة أخرى تسمى بالفرضيّة البديلة ونرمز لها بالرمز H_1 ، في حالة متوسط أطوال الذكور البالغين يمكن أن يأخذ الفرض البديل إحدى الحالات التالية:

$\mu \neq 165$ أو $\mu < 165$ أو $\mu > 165$ أو $\mu = 170$ ... الخ

(١١ - ٩) اختبارات المعنويـة

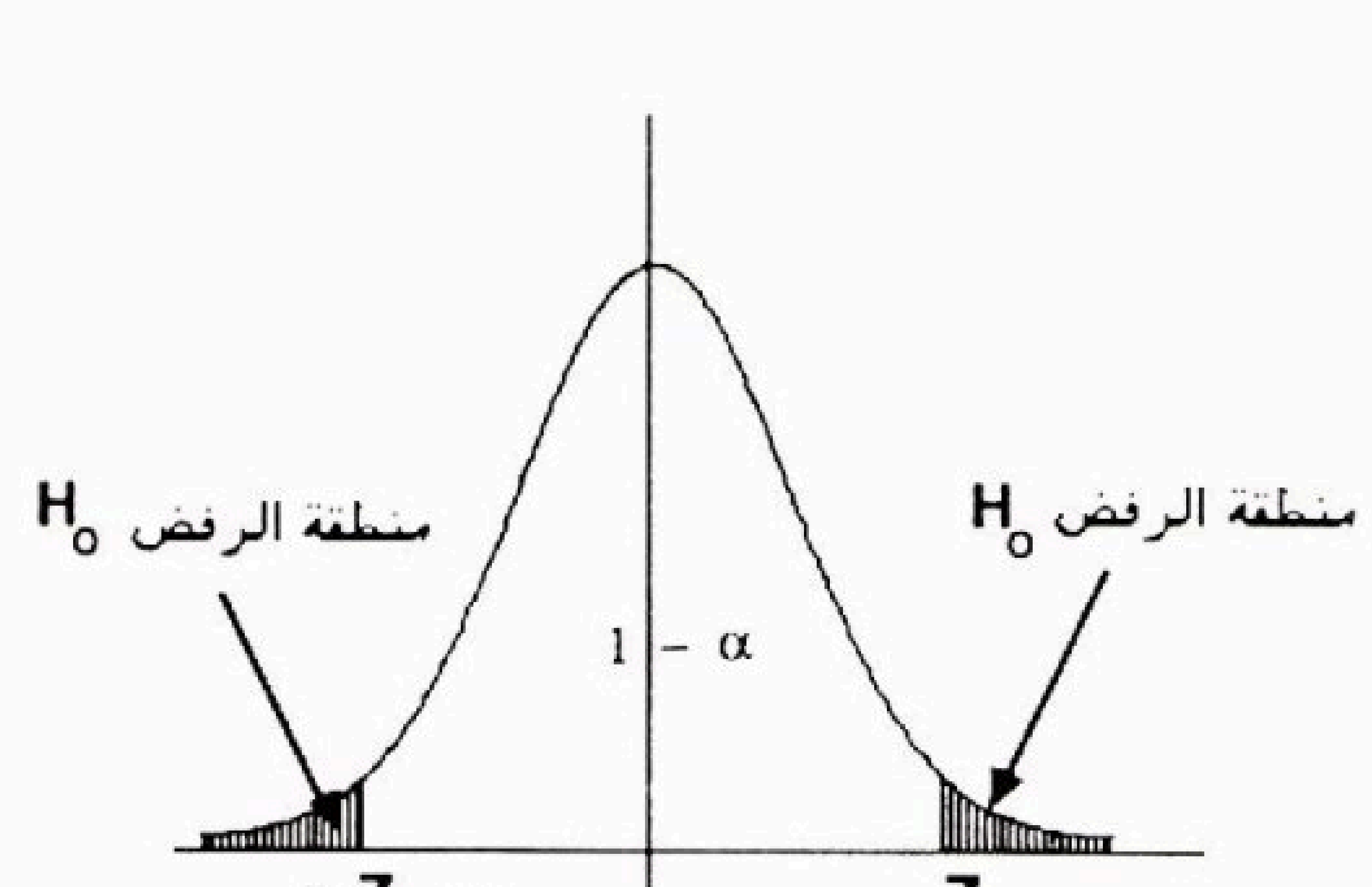
ولاختبار صحة الفرضيّة الأولية H_0 يجب علينا تكوين إحصاءة وهي دالة من مشاهدات العيّنة العشوائية ومعالم H_0 مثل $Z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}}$ ، وعادة يكون توزيع الإحصاءة

معروفًا ويقسم المجال المقابل لهذه الإحصاءة إلى منطقتين، المنطقة الأولى يمكن تسميتها منطقة القبول وهي التي يكـون فيها احتمال حدوث قيم الإحصاءة $(1 - \alpha)$ كبيرًا عندما تكون الفرضيّة الأولية H_0 صحيحة. والمنطقة الثانية تسمى بمنطقة الرفض وهي التي يكون احتمال حدوث قيم الإحصاءة α صغيرًا عندما تكون الفـرضية الأولية H_0 صحيحـة. وفي حالـة متـوسط أطـوال الـذكـور البـالغين نأخـذ الإحصـائية $Z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}}$ حيث سبق معرفة توزيعها بأنه يقترب من التوزيع الطبيعي القياسي $N(0,1)$ ويمكن توضيح منـطقة القبول والرفض للفرضيّة $H_0 : \mu = 160$ على أساس صحة الفرضيّة H_0 مع الأخذ في الاعتبار الفرضيّة البديلة H_1 في الحالات التالية:

١) $H_0 : \mu = 160$,

$H_1 : \mu \neq 160$,

في هذه الحالة الفرض البديل H_1 له طرفان كما هو موضح بالمنطقة المظللة في الرسم التالي شكل (١١ ـ ٢).



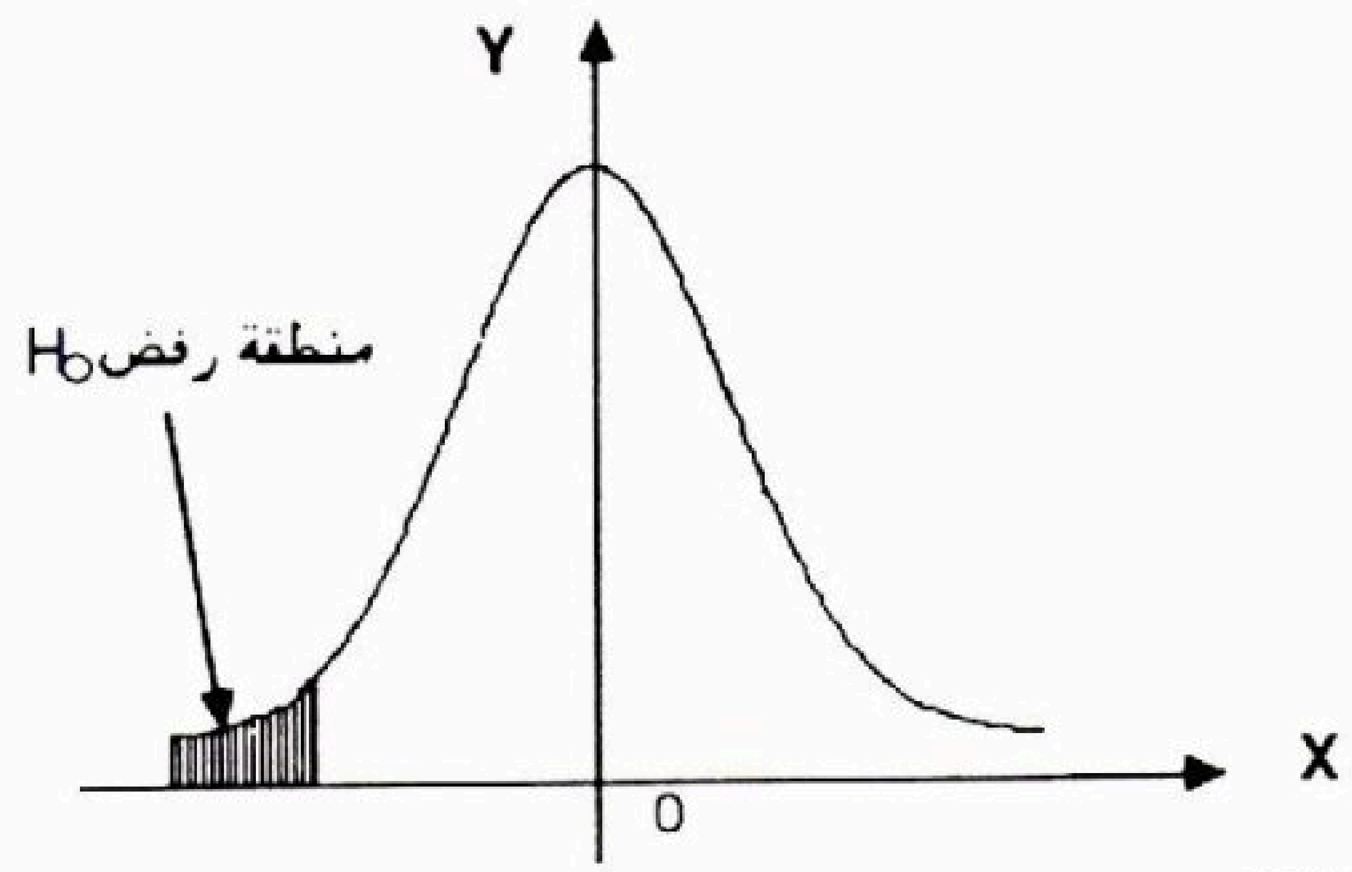
شكل (١١ ـ ٢) يمثل منطقة الرفض للفرض H_0 من الطرفين

منطقة قبول H_0 هي $-z_{\alpha/2} \leq Z \leq z_{\alpha/2}$. وهي غير مظللة في شكل (١١ ـ ٢) السابق.

٢) $H_0 : \mu = 160$ ،

$H_1 : \mu < 160$

الفرض البديل له طرف واحد أدنى ونوضح ذلك بالشكل (١١ - ٣) التالي:



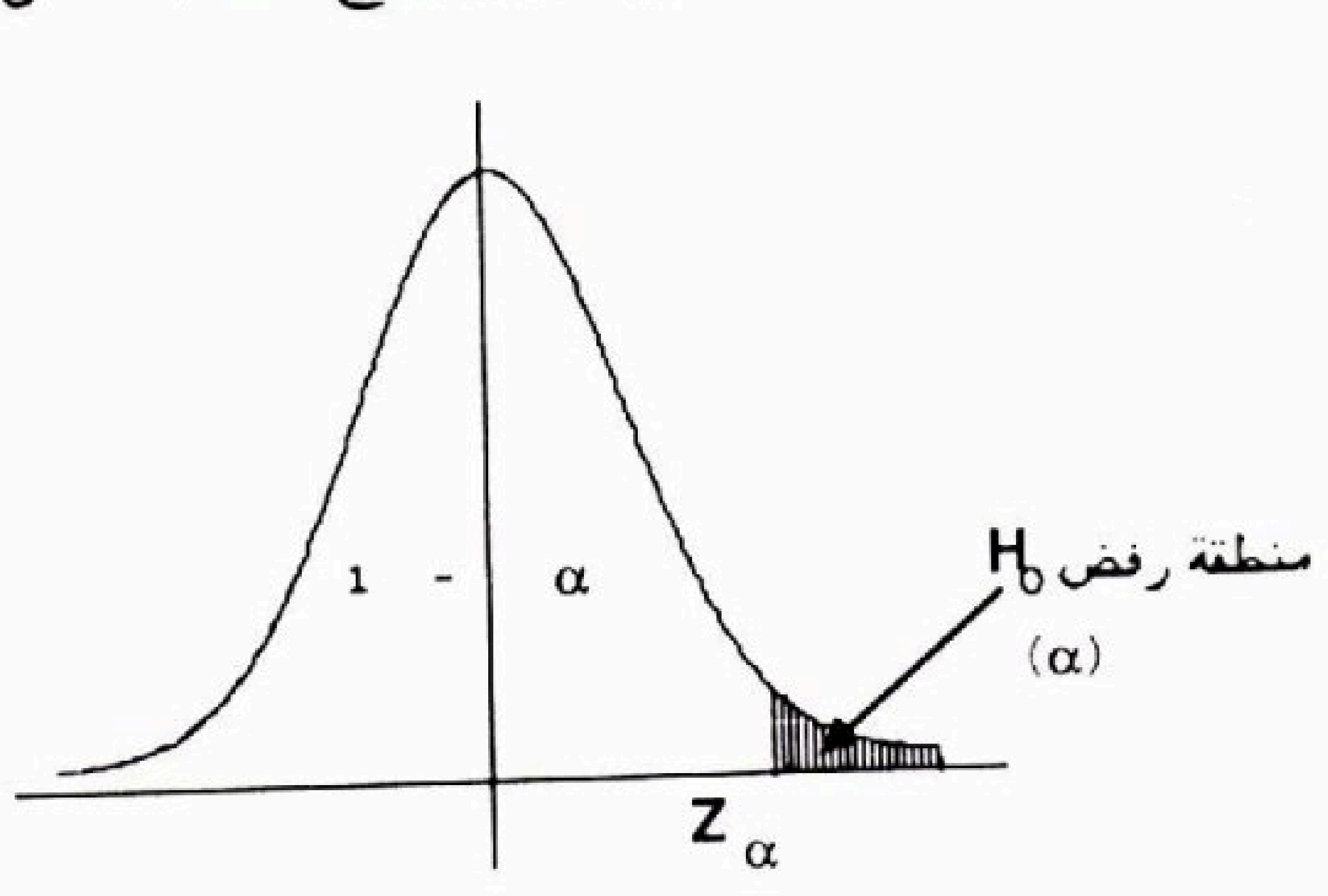
شكل (١١ - ٣) يمثل منطقة الرفض للفرض H_0 من الطرف الأيسر

منطقة قبول H_0 هي $Z > -z_\alpha$.

٣) $H_0 : \mu = 160$ ،

$H_1 : \mu > 160$

الفرض البديل له طرف واحد من أعلى ونوضح ذلك بالشكل (١١ - ٤) التالي:



شكل (١١ - ٤) يمثل منطقة الرفض للفرض H_0 من الطرف الأيمن

منطقة قبول H_0 هي $Z < z_\alpha$.

وعمومًا إن كانت الفرضيّة الأولية $H_0 : \mu = \mu_0$ فإنه يمكن تلخيص مناطق الرفض والقبول للفرضيّة الأولية وذلك حسب الفرض البديل H_1 . وذلك لقيم الإحصاءة z المحسوبة من العيّنة كالتالي :

الفرض البديل H_1	$\mu \neq \mu_0$	$\mu < \mu_0$	$\mu > \mu_0$
منطقة الرفض لـ H_0	$Z < -z_{\alpha/2}$ أو $Z > z_{\alpha/2}$	$Z < -z_\alpha$	$Z > z_\alpha$
منطقة القبول لـ H_0	$-z_{\alpha/2} < Z < z_{\alpha/2}$	$Z > -z_\alpha$	$Z < z_\alpha$

(١١ ـ ١٠) الخطأ من النوع الأول α والخطأ من النوع الثاني β

أي قرار إحصائي يترتب عليه نوعان من الخطأ. خطأ من النوع الأول: هو عندما نرفض الفرضيّة الأولية H_0 وهي صحيحة، ويكون احتمال هذا الخطأ هو قيمة α ، وهي عادة ما تكون صغيرة ومثال على ذلك .... ,0.01 ,0.05 ,0.1 وأحيانًا تسمى α مستوى المعنوية. وعندما نقبل الفرضيّة الأوليّة H_0 وهي خاطئة نكون قد ارتكبنا خطأ من النوع الثاني، ويكون احتمال هذا الخطأ هو القيمة β . ويمكن تلخيص ذلك في الجدول التالي :

القرار / الفرضيّة H_0	قبول H_0	رفض H_0
الفرضية H_0 صحيحة	قرار صحيح	خطأ من النوع الأول α
الفرضيّة H_0 خطأ	خطأ من النوع الثاني β	قرار صحيح

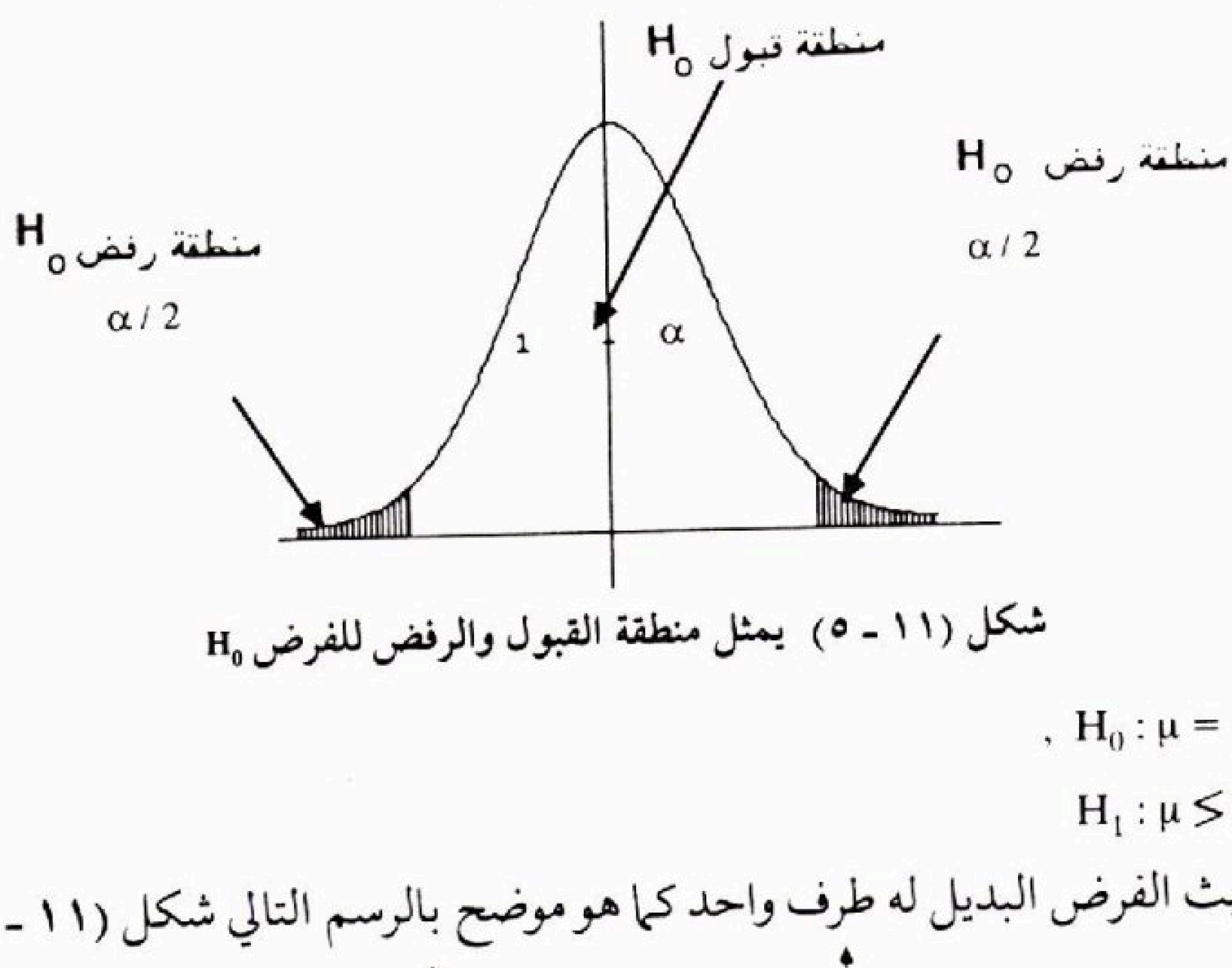
(١١ - ١١) اختبار الفرضيّات للمتوسط μ

عندما يكون الانحراف المعياري σ للمجتمع معلومًا فإن الإحصائية $\frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$ يكون لها توزيع يقترب من التوزيع الطبيعي القياسي $N(0,1)$ وبذلك يمكننا من تكوين الفرضيّة الأولية (فرض العدم) H_0 والفرضيّة البديلة H_1 كما يلي :

١) $H_0 : \mu = \mu_0$ ،

$H_1 : \mu \neq \mu_0$

حيث الفرض البديل له طرفان كما هو موضح بالرسم التالي شكل (١١ - ٥) :

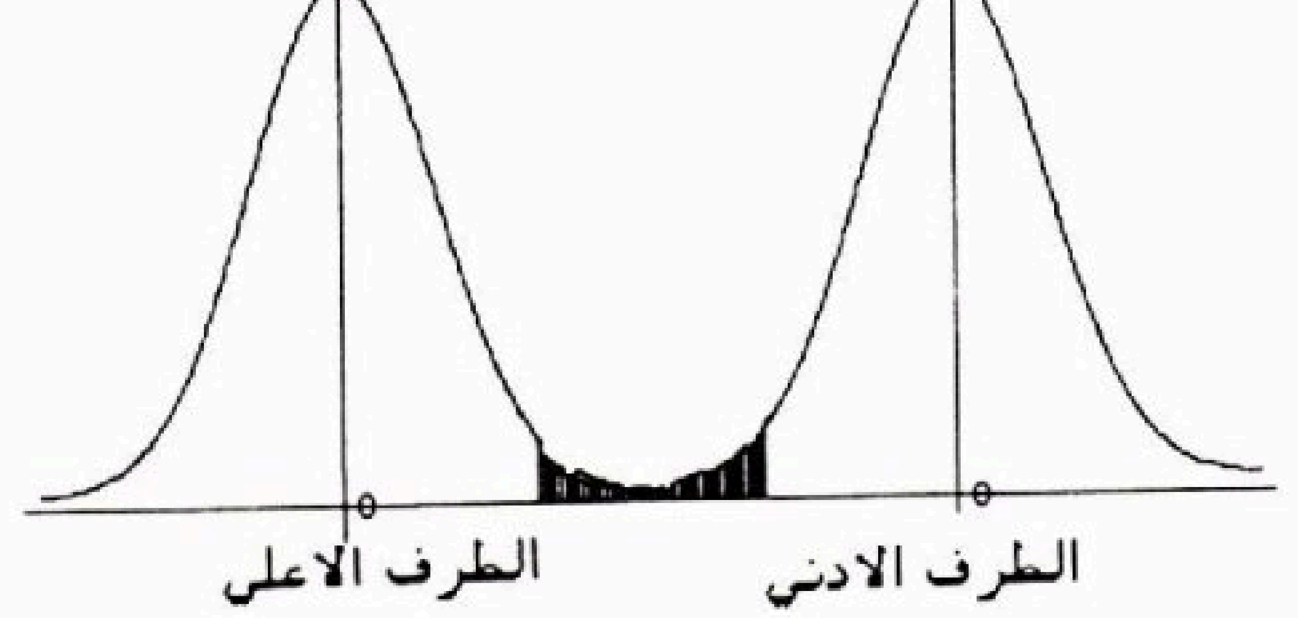


شكل (١١ - ٥) يمثل منطقة القبول والرفض للفرض H_0

٢) $H_0 : \mu = \mu_0$ ،

$H_1 : \mu \lessgtr \mu_0$

حيث الفرض البديل له طرف واحد كما هو موضح بالرسم التالي شكل (١١ - ٦) :



شكل (١١ - ٦) يمثل منطقة الرفض للفرض H_0 للطرف الأدنى والطرف الأعلى

ثم نكون الإحصائية ($Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$) ومنها نحسب قيمة Z باستخدام البيانات المشـاهدة من العيّنة ثم نختار مستوى المعنوية α الذي على أساسه نحدد القيم الحـرجة $Z_{\alpha/2}$, $Z_{\alpha/2}$ - عندما تكون الفرضيّة البديلة H_1 من طرفين ونوضح ذلك بالمثال التالي :

مثال (١١ - ٩)

أخذت عيّنة مكونة من 64 ذكر بالغ في بلد ما فوجد أن متوسط الطول لهم هو 155 سم وكان الانحراف المعياري للمجتمع معلومًا ويساوي 5 سم . اختبر الفرض القائل أن متوسط هذا المجتمع $\mu = 160$ عند مستوى معنوية (ا) 0.05 ; (ب) 0.01 .

الحـــل

نكون الفرضيّة الأولية والبديلة كالتالي :

$H_0 : \mu = 160$ cm ,

الفرض البديل من الطرف الأدنى : $H_1 : \mu < 160$ cm

نحسب الإحصائية التالية :

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

حـيث $\bar{x} = 155$ cm , $\mu_0 = 260$

$\sigma = 5$, $n = 64$

فيكون z_0 المحسوبة كالتالي :

$$Z_0 = \frac{150 - 160}{\frac{5}{\sqrt{64}}} = -8$$

(ا) عند مستوى معنوية $\alpha = 0.05$ فإن قيمة $z_{0.05}$ من الجدول هي

$$z_{0.05} = -1.64$$

أي أن

$$Z_0 < z_{0.05}$$

أي أن قيمة Z واقعة في منطقة الرفض لـ H_0 وبذلك نرفض H_0 عند مستوى معنوية 0.05

(ب) عند مستوى معنوية $\alpha = 0.01$ فإن قيمة $z_{0.01}$ من الجدول هي :

$$z_{0.01} = -2.33$$

أي أن :

$$Z_0 < z_{0.01}$$

أي Z_0 واقعة في منطقة الرفض لـ H_0 وبذلك نرفض H_0 عند مستوى معنوية 0.01 . أي أن قيمة $z_0 = -8$ المحسوبة من العيّنة عالية المعنوية والعيّنة لا تنتمي لمجتمع متوسطه 160 سم .

ويمكن تلخيص بعض القيم الحرجة z عند مستويات معنوية مختلفة في الجدول التالي :

مستوى المعنوية α	0.1	0.05	0.01	0.005	0.002
قيم z الحرجة للفرض البديل من طرف واحد	±1.28	±1.645	±2.33	±2.58	±2.88
قيم z الحرجة للفرض البديل من طرفين	±1.645	±1.96	±2.58	±2.81	±3.08

(١١ ـ ١١ ـ ١) درجات الحرية

إن كان لدينا مجتمع ما ويوجد به عدد من المعالم. ونرغب في تقدير هذه المعالم من عيّنة تحتوي على n من البيانات المستقلة. فإن درجة الحرية التي يرمز لها بالرمز ν

تساوي عدد البيانات المستقلة للعيّنة مطروحًا منها عدد المعالم للمجتمع التي يجب تقديرها من العينة ويعبر عن ذلك رياضيًّا كمايلي :

$$v = n - K$$

حيث K تساوي عدد المعالم التي يجب تقديرها . فإن كان المطلوب تقدير المتوسط μ بـ $\bar{x}$ من العيّنة فإن $K = 1$ وبالتالي $(v = n - 1)$ وإن كان المطلوب تقدير كل من μ و σ فإن $K = 2$ أي أن درجات الحرية في هذه الحالة $(v = n - 2)$ وهكذا . . . .

(١١ ـ ١١ ـ ٢) اختبار الفرضيّات للمتوسط μ (العيّنات الصغيرة المأخوذة من المجتمع الطبيعي)

وذلك عندما يكون الانحراف المعياري σ مجهولاً وحجم العيّنة صغيرًا $(n<30)$ فإننا نستبدل الإحصائية Z بالإحصائية $t = (\bar{x} - \mu_0) / (S/\sqrt{n})$ التي يكون لها توزيع معاينة يتبع توزيع t بدرجات حرية قدرها $(v = n - 1)$ والقيم الحرجة عند درجات حرية هي $t_{\alpha/2}$ و t_α بدلاً من $z_\alpha, z_{\alpha/2}$ على الترتيب . وذلك للإحصائية السابقة Z وتوضح طريقة الاختيار للفرضيّة الأولية H_0 في هذه الحالة بالمثال التالي :

مثال (١١ ـ ١٠)

اختبرت 9 حبال من إنتاج المصانع لمعرفة قوة مقاومتها للقَطْع . فأظهرت متوسط مقاومة للقَطْع هو 6750 ث . كجم بانحراف معياري (s) يساوي 240 ث . كجم بينما يدّعي المصنع المنتج الرقم 7000 ث . كجم كقوة للقَطْع للإنتاج . هل يمكن تأييد إدّعاء المصنع عند مستوى معنوية (ا) 0.05 ، (ب) 0.01 .

الحــل

نكوّن اختبار الفرضيّات كالتالي :

$$H_0 : \mu_0 = 7000 \quad , \quad H_1 : \mu < 7000$$

نجد أنه اختبار من طرف واحد .

نحسب الإحصائية :

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

حيث

$$\bar{x} = 6750 \text{ Kg}, \quad S = 240 \text{ Kg}, \quad n = 9$$

أي قيمة t المحسوبة هي :

حيث
$$t = \frac{6750 - 7000}{240/\sqrt{9}}$$

$$= \frac{-250 \times 3}{240} = \frac{-75}{24} = -3.125$$

ا) الاختبار عند مستوى معنوية 0.05 أي أن قيمة $t(0.05,8)$ من الجدول هي :

$$t(0.05,8) = -1.860$$

أي أن

$$t(0.05,8) > t$$

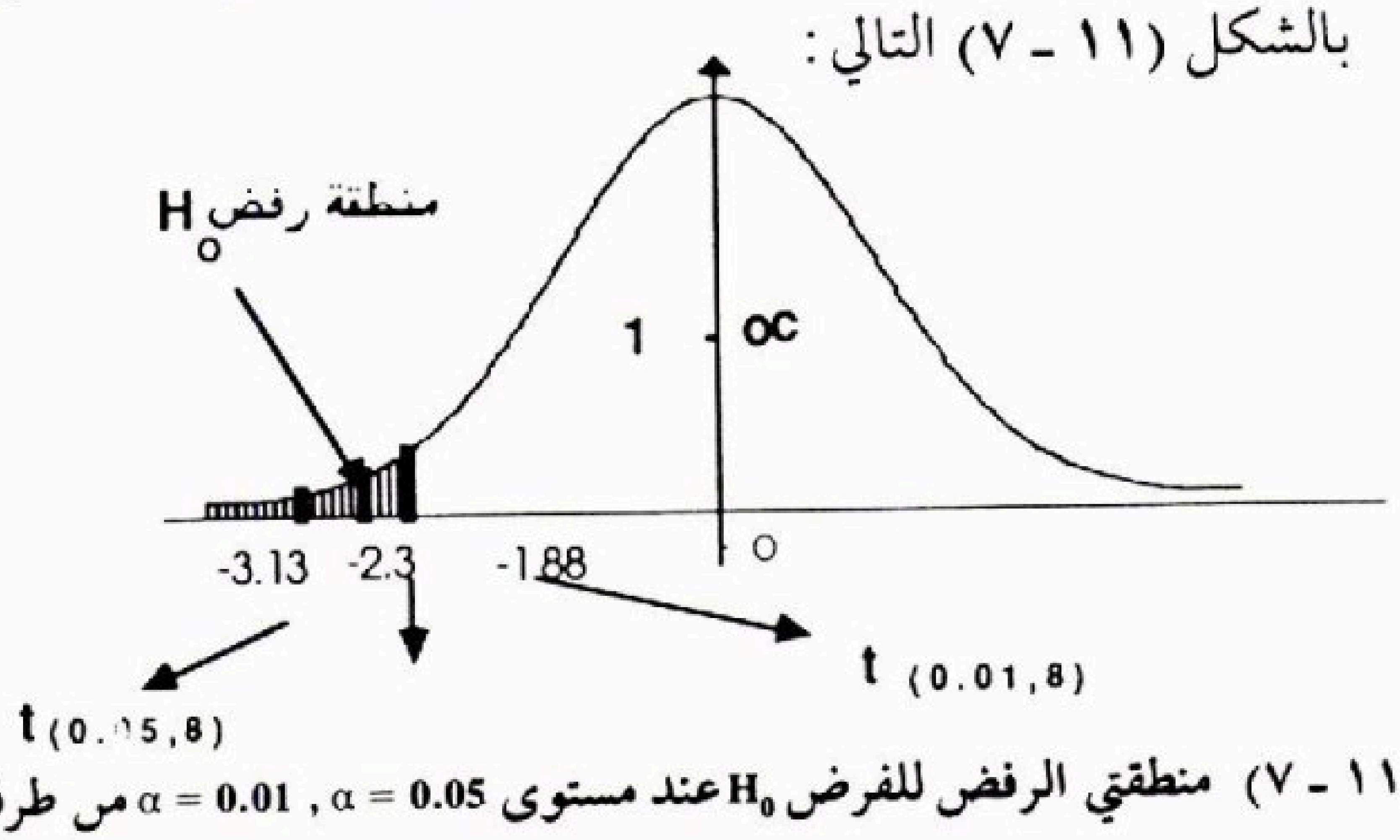
فإننا نجد أن t المحسوبة واقعة في منطقة الرفض لـ H_0 .

ب) الاختبار عند مستوى معنوية 0.01 .

$$t = (0.01,8) = -2.306$$

$$t(0.01,8) > t$$

أي أن t المحسوبة واقعة في منطقة الرفض لـ H_0 ونوضح ذلك بالشكل (١١ ـ ٧) التالي :



شكل (١١ ـ ٧) منطقتي الرفض للفرض H_0 عند مستوى $\alpha = 0.05$, $\alpha = 0.01$ من طرف واحد

t المحسوبة واقعة في منطقة الرفض لـ H_0 لكل من مستوى المعنوية 0.05 و 0.01 أي أن أي أن ادّعاء المصنع غير صحيح.

(١١ ـ ١١ ـ ٣) اختبار الفرضيّات للفرق بين المتوسطات للعيّنات الكبيرة

توجد كثير من المسائل الإحصائية يكون المطلوب أن نقرر ما إذا كان الفرق بين متـوسط عيّنتـين مستقلتين يرجع إلى عامل الصدفة فقط أم هناك فروقًا حقيقية مثل متوسط أطوال عيّنتين من الرجال من مجتمعين مختلفين ومتوسط إنتاج القمح في منطقتين مختلفتين ومتوسط النجاح لطريقتين للتدريس و. . . . إلخ . ولقد وجد الإحصائيون أنه إن كان المتوسطان $\bar{x}_2$ و $\bar{x}_1$ من عينتين عشوائيتين حجمهما على الترتيب n_2 و n_1 فإن توزيع المعاينة للفرق $\bar{X}_1 - \bar{X}_2$ يقترب من التوزيع الطبيعي بمتوسط $(\mu_1 - \mu_2)$ وانحراف معياري $\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ حيث (μ_1, μ_2) ، (σ_1, σ_2) هما متوسطا وانحرافا المجتمعين اللذين أخذنا منهما العيّنتين. وعادة في الحياة العملية تكون σ_2 و σ_1 مجهولتين، لذا فإنه من الممكن أن نستعيض عنهما بالانحراف المعياري للعيّنتين S_2 و S_1 كتقدير لـ σ_2 و σ_1 بحيث يكون حجم كل من العيّنتين كبيراً أي $(n_2 > 30)$، $(n_1 > 30)$ ويمكن تكوين الفرضية الأوليّة كما يلي:

$$H_0 : \mu_1 - \mu_2 = 0$$

والفرضيّة البديلة H_1 أحد الفرضيّات التالية:

$$H_1 : \mu_1 - \mu_2 \lessgtr 0 \quad , \quad H_1 : \mu_1 - \mu_2 \neq 0$$

ولتوضيح ذلك نأخذ المثال التالي.

مثال (١١ ـ ١١)

إن كنا نريد اختبار الفرض القائل إنه لا توجد فروق معنوية بين أطوال الذكور البالغين المولودين في قطرين مختلفين حيث كانت نتائج المعاينة ملخصة كمايلي:

$$n_1 = 120 \quad , \quad \bar{x}_1 = 62.7 \quad m \quad S_1 = 2.50$$

$$n_2 = 150 \quad , \quad \bar{x}_2 = 61.8 \quad m \quad S_2 = 2.62$$

حيث كانت القياسات بالبوصة .

ننشىء اختبارات الفرضيّات الأولية والبديلة كالتالي :

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

ثم نحدد مستوى المعنوية $\alpha = 0.01$ وبالتالي :

نحسب الإحصائية Z_0 كالتالي :

$$Z_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{62.7 - 61.8}{\sqrt{\frac{(2.50)^2}{120} + \frac{(2.62)^2}{150}}} = 2.88$$

تكون قيمة $Z_{0.01}$ من الجدول الطبيعي القياسي هي :

$$Z_{0.01} = 2.58$$

$$Z_0 = 2.88 > Z_{0.01} = 2.58$$

أي قيمة Z_0 المحسوبة واقعة في منطقة رفض H_0 وبذلك نرفض الفرضيّة الأولية التي تقول إنه لا يوجد اختلاف في أطوال الذكور البالغين في كل من القطرين بمستوى معنوية 0.01 ؛ أي بدرجة ثقة قدرها 99% .

(١١ ـ ١١ ـ ٤) اختبار الفرضيّات بين المتوسطات في العيّنات الصغيرة المأخوذة من المجتمع الطبيعي

لقد وجد الإحصائيون أن الفرق بين المتوسطين $\bar{X}_1, \bar{X}_2$ عندما يكون حجم العيّنتين المستقلتين $n_1 ; n_2$ صغيراً . وأن الانحراف المعياري لكل من المجتمعين المأخوذ

منهما العيّنتان هما $\sigma_1 = \sigma_2$ المجهولتين فإن الإحصائية التالية :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

حيث

$$s^2 = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2}$$

تتبع توزيع t بدرجات حرية قدرها $v = n_1 + n_2 - 2$ ونوضّح طريقة اختبار الفرق بين متوسطي عيّنتين مستقلتين بالمثال التالي .

مثال (١١ ـ ١٢)

في دراسة للمقارنة بين السعرات الحرارية الناتجة لنوعين من الفحم المنتج من منجمين مختلفين كانت النتائج التالية بملايين السعرات الحرارية هي :

المنجم الأول : 7930, 7860, 8380, 8230, 8400

المنجم الثاني : 7660, 8070, 7720, 7690, 7510

اختبر الفرض القائل أن المنجمين لها السعرات الحرارية نفسها عند مستوى معنوية $\alpha = 0.05$.

الحـــل

نحسب القيم التالية :

$$\bar{x}_1 = 8160 \quad , \quad \bar{x}_2 = 7730$$

$$s_1 = 251.89 \quad , \quad s_2 = 206.519$$

$$n_1 = 5 \quad , \quad n_2 = 5$$

ثم نكوّن اختبار الفرضيات كالتالي :

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

ثم نحسب الإحصائية t_0 كالتالي :

$$t_{\triangle} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2 (\frac{1}{n_1} + \frac{1}{n_2})}}$$

حيث

$$s^2 = \frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}$$

$$\therefore t_0 = \frac{8160 - 7730}{\sqrt{\frac{253800 + 170600}{5 + 5 - 2} (\frac{1}{5} + \frac{1}{5})}} = 2.95$$

$$v = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$$

ونختار مستوى المعنوية $\alpha = 0.05$ فيكون :

$$t_{(0.025,8)} = 2.306$$

فتكون قيمة

$$t_{(0.025,8)} < t_0$$

أي قيمة t المحسوبة واقعة في منطقة الرفض H_0 أي أن السعرات الحرارية لكل من النوعين مختلفة .

(١١ ـ ١٢) اختبار الفرضيّات للنسبة R

إن كان المجتمع الذي نقوم بدراسته يتوزع توزيعًا متقطعًا مثل التوزيع ذي الحدين ـ وإن فرضنا أن نسبة ظاهرة معيّنة في هذا المجتمع R وكانت r هي نسبة هذه الظاهرة في عيّنة حجمها n $(n \geq 30)$ ـ فإن المطلوب معرفة وجود فرق معنوي بين نسبة الظاهرة في كل من العيّنة والمجتمع . ولقد سبق أن علمنا بأن الإحصائية $\frac{r - R}{\sqrt{rq/n}}$ تقترب من التوزيع الطبيعي القياسي ، بذلك يمكننا تكوين الفرضيّة الأولى H_0 ، والفرضيّة البديلة H_1 بالطرق التالية :

i) $H_0 : R = R_0$.

$H_1 : R \neq R_0$ حيث الفرض البديل له طرفان .

(ii $H_0 : R = R_0$,

$H_1 : R \lesseqgtr R_0$ حيث الفرض البديل له طرف واحد.

وبعد تكوين الفرضيّة الأولية H_0 والفرضيّة البديلة H_1 ثم ننشىء الإحصائيّة $Z = \frac{r - R_0}{\sqrt{R_0 Q_0 / n}}$ ونحسب قيمة Z باستخدام البيانات المشاهدة من العيّنة والفرضيّة H_0 ثم نختار مستوى العينة α وهو الذي على أساسه نحدد القيم الحرجة Z_α, Z_α - عندما تكون الفرضيّة البديلة H_1 لها طرف واحد، أو $Z_{\alpha/2}, Z_{\alpha/2}$ - عندما تكون الفرضيّة البديلة H_1 من طرفين، ونوضّح ذلك بالمثال التالي.

مثال (١١ - ١٣)

مصنع للأدوية المسجلة يدّعي أن دواء من إنتاجه له فاعلية بنسبة 90% في التخفيف من الحساسية لفترة 8 ساعات في عيّنة من 200 شخص مصابين بالحساسية، أدى الدواء إلى تخفيف آلام 160 منهم. قرر ادّعاء المصنع صحيح أم غير صحيح عند مستوى معنوية $\alpha = 0.01$.

الحـــل

نكوّن الفرضيّة الأولية والبديلة كالتالي:

$H_0 : R = 0.9$ (الإدّعاء صحيح).

$H_1 : R < 0.9$ الفرض البديل من الطرف الأدنى (الإدعاء باطل).

نحسب r وهي: $r = \frac{160}{200} = 0.8$

وتكون Z_0 المحسوبة كالتالي:

$$Z_0 = \frac{r - R_0}{\sqrt{R_0 Q_0 / n}} = \frac{0.8 - 0.9}{\sqrt{0.9 \times 0.1 / 200}} = -4.73$$

عندما يكون مستوى المعنوية $\alpha = 0.01$ فإن قيمة Z الجدولية تساوي $Z_{(0.01)} = -2.33$. وعليه فإن $Z_0 < Z_{0.01}$ المحسوبة وبذلك يكون الاختبار معنويًا أي أن الادّعاء غير صحيح.

(١١ ـ ١٢ ـ ١) اختبار الفرضيّات للفرق بين النسبتين $(R_1 - R_2)$

اختبار الفرق بين نسبتين في عيّنتين لمعرفة ما إذا كان الفرق يرجع إلى عامل الصدفة أم أن هناك فروقًا حقيقية؟ مثل نسبة التدخين لعيّنتين من الطلاب في جامعتين مختلفتين ونسبة الشفاء من مرض معيّن باستخدام عقارين مختلفين لعيّنتين من المرض. . . إلخ . فإذا كانت النسبتين r_2 و r_1 من عيّنتين عشوائيتين حجمهما على الترتيب n_2 و n_1 فإن توزيع المعاينة للفرق $r_1 - r_2$ يقترب من التوزيع الطبيعي بتوقع $(R_1 - R_2)$ وانحراف معياري

$$\sqrt{\frac{R_1(1-R_1)}{n_1} + \frac{R_2(1-R_2)}{n_2}}$$

ويمكن تكوين الفرضيّة الأولية كالتالي :

$$H_0 ; R_1 - R_2 = 0$$

والفرضيّة البديلة وهي :

$H_1 : R_1 - R_2 \neq 0$ الفرضيّة البديلة من طرفين .

أو

$H_1 : R_1 - R_2 \lessgtr 0$ الفرضيّة البديلة من طرف واحد .

ونوضّح ذلك بالمثال التالي .

مثال (١١ ـ ١٤)

مجموعتان B و A ، تتكون كل منهما من 100 شخص مصابين بمرض معيّن . أعطي مصل للمجموعة A ولم يعط للمجموعة B (التي تسمى بالمجموعة الضابطة)، بخلاف ذلك فإن المجموعتين تعاملان معاملة متماثلة . وقد وجد أنه في المجموعة A شفي 75 شخصًا من المرض، بينما في المجموعة B شفي 65 شخصًا . أختبر الفرض أن المصل يساعد على الشفاء من المرض باستخدام مستوى معنوية 0.05 .

الحـل

$r_1 =$ النسبة في المجموعة A للأشخاص الذين شفوا .

$r_2 =$ النسبة في المجموعة B للأشخاص الذين شفوا .

$$\because\ r_1 = 0.75,\ r_2 = 0.65$$

$$\therefore\ r = \frac{r_1 n_1 + r_2 n_2}{n_1 + n_2} = \frac{75 + 65}{200} = 0.70$$

نكوّن الفرضية الأولية والبديلة كالتالي :

$H_0 : R_1 - R_2 = 0$ المصل غير فعال .

$H_1 : R_1 - R_2 > 0$ المصل فعال .

نحسب z كالتالي :

$$Z_0 = \frac{r_1 - r_2 - 0}{\sqrt{rq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.75 - 0.65}{\sqrt{0.7 \times 0.3\left(\frac{1}{100} + \frac{1}{100}\right)}} = 1.54$$

قيم $Z_{(0.05)}$ من الجدول هي : $Z_{(0.05)} = 1.645$

فإننا نلاحظ أن قيم $Z_{0.05} > Z_0$ المحسوبة .

ولـذلـك تكـون النتائج راجعة للصدفة عند هذا المستوى وبالتالي لا نستطيع رفض الفرضيّة الأولية H_0 والمصل غير فعّال .

(١١ ـ ١٣) تمـــارين

(١) أوجـد حجم الخط الأقصى عنـد استخـدام المتـوسط $\bar{X}$ والذي أخذ من عيّنة حجمها 40 علمًا بأن الانحراف المعياري لتلك العيّنة هو 1.45 ومستوى العيّنة كان 0.95 .

(٢) في السؤال (١) أوجد فترة الثقة عند مستوى المعنوية $\alpha = 0.05$, $\alpha = 0.01$ وطول كل منها ثم قارن بينها .

(٣) لدراسة النمو لنوع خاص من الزهور أخذت عيّنة من 50 زهرة وجد أن متوسط النمو خلال العام هو 44.8 سم والانحراف المعياري هو 4.7 . أوجد حدود الثقة

للمتوسط الحقيقي للنمو عند درجة ثقة :

(i) 95 % (ii) 99 %

(٤) في السؤال (٣) السابق أوجد الحجم الأقصى للخطأ في التقدير باحتمال قدره 0.99

(٥) في مسح اجتماعي اشتمل على 300 عائلة وجد أن متوسط ما أنفق على الطعام خلال عام هو 3943 ريالاً بانحراف معياري قدره 415 ريالاً .
أوجـد :
i) فترة الثقة باحتمال قدره 0.99 .
ii) ماذا يمكن القول عند الاحتمال 0.95 عن الحجم الأقصى للخطأ .

(٦) لدراسة متوسط عدد الساعات التي يقضيها الطلاّب في مشاهدة التلفزيون خلال أسبوع . ما هو حجم العيّنة المطلوب للدراسة إن وجدنا أنه من الضروري أخذ انحراف معياري قدره 3.2 ساعة حتى نحصل على حقيقة أن متوسط العيّنة هذه يختلف عن متوسط المجتمع بأقل من 24 دقيقة باحتمال 0.95 .

(٧) ما هو حجم العيّنـة المطلوب لنتوصل إلى حقيقة أن متوسط العيّنة يختلف عن متـوسط المجتمـع بأقل من 45 وذلك باحتمال 0.99 إن كان الانحراف المعياري يساوي 320 .

(٨) في اختبار للزمن الذي يستغرقه تجميع ماكينة معيّنة وجد أن الزمن الذي استغرقته 6 ماكينات هو على التوالي : 11 ,12 ,6 ,12 ,14 ,13 دقيقة .
أوجد فترة الثقة للمتوسط عن مستوى معنوية 0.05 .

(٩) أخذت العيّنة 1.7 ,2.3 ,0.7 ,1.1 ,1.0 من توزيع طبيعي $N(\mu, 1)$.
أوجد فترة 95% وفترة 99% ثقة للمتوسط μ .

(١٠) أخذت عيّنة من 25 طالبًا وقيست أوزانهم فوجد أن وسطها 6.3 كجم وانحرافها المعياري 9 كجم . أوجد :

i) 90% فترة ثقة لمتوسط الأوزان .

ii) 95% فترة ثقة لمتوسط الأوزان .

(١١) أراد مدير مصنع للإسمنت إيجاد فترة 95% ثقة لمتوسط وزن كيس الإسمنت الذي ينتجه المصنع . فإذا كان هذا المدير يعلم أن الانحراف المعياري لوزن أكياس الإسمنت يساوي 1.2 كجم فما حجم العيّنة من الأكياس التي يجب أن تجرى عليها التجربة حتى لا يزيد طول فترة الثقة المطلوبة عن 2.6 كجم .

(١٢) عيّنة من 10 قياسات لأقطار كرة أعطيت متوسطًا 4.38 ملم وانحرافًا معياريًا 0.06 ملم .

أوجـد :

(أ) 95 % (ب) 99 % حدود ثقة للقطر الفعلي .

(١٣) لقياس زمن رد الفعل ، قدّر عالم سيكلوجي الانحراف المعياري بـ 0.05 ثانية . ما هو حجم العيّنة من القياسات بحيث يكون (i) 95 % ,(ii) 99 % واثقين أن الخطأ تقديره لن يتجاوز 0.01 ثانية؟

(١٤) باستخدام النتائج الآتية وباحتمال 95 % حدد أيًا من العيّنات التالية مسحوبة من مجتمع متوسطه 800 وانحرافه المعياري 50 .

العيّنـة	حجم العيّنة	المتوسـط
1	100	805
2	200	805
3	400	805

ثم أحسب حدود الثقة للمتوسط في كل عيّنة مع احتمال 99 % .

(١٥) مصنع لإنتاج إطارات السيارات يدّعي أن إنتاجه من الإطارات يصلح للاستعمال لمسافة 300,000 كم، عند استخدام نوع معينّ من السيارات أخذت عيّنة عشوائية وكانت النتائج كالتالي:

الإطار	1	2	3	4	5	6	7	8
المسافة	31,000	26,000	28,000	25,000	32,000	29,000	32,000	23,000

والمطلوب وباحتمال % 95 اختبار الادّعاء صحيح أم لا. ثم أحسب حدود الثقة لمتوسط المسافة التي يصلح لها الإطار.

(١٦) لمعرفة أثر غذاء معينّ على زيادة الوزن أخذت عيّنة من خمسة فئران وتم تغذيتها بهذا الغذاء وكانت أوزانها بعد التغذية هي:

2.4, 2.3, 1.5, 1.4, 1.6

أفنستطيع أن نحكم على أن هذه العيّنة مأخوذة من مجتمع متوسط الوزن منه 1.8 أم لا وذلك باحتمال % 95 .

(١٧) من البيانات الآتية واحتمال % 99 اختبر الفرض القائل أن متوسط كل المجتمعين المأخوذة منه هذه البيانات متساوية.

	الأولى	الثانية
الانحراف المعياري	1.2	1.6
حجم العيّنة	50	50
المتوسط	22.3	21.6

ثم أحسب حدود الثقة للفرق بين المتوسطين

(١٨) اختيرت مجموعتان من الأرانب، الأولى من 12 أرنبًا أعطيت الغذاء (ا) والثانية من 15 أرنبًا أعطيت الغذاء (ب) وكانت الزيادة في الوزن بعد فترة معيّنة هي :

(ا) 35, 31, 30, 24, 22, 13, 25, 24, 34, 28, 30, 26

(ب) 21, 18, 35, 32, 30, 40, 31, 47, 8, 22, 34, 44, 22, 27, 35

اختبر معنوية الفرق بين أثري الغذائين عند درجة ثقة 95 % .

(١٩) أخذت عيّنة حجمها n = 4 وأعطت القيم التالية :

14.29, 14.33, 12.27, 14.31

ماذا يمكن أن يقال باحتمال 0.99 عن الحجم الأقصى للخطأ إن استعمل متوسط تلك العينة للأغراض الخاصة.

(٢٠) طبيب أسنان وجد أن 6 من مرضاه يحتاجون إلى 3 ,4 ,0 ,6 ,3 ,2 عمليات حشو.

i) ما حجم الخطأ إن استخدم هذا الطبيب متوسط هذه العيّنة لتقدير متوسط المجتمع عند مستوى معنوية 0.01 .

ii) أوجد فترة الثقة عند مستوى المعنوية 0.1 .

(٢١) في حالة قضائية حيث هناك فرد يحاكم في تهمة سرقة أي نوعيْ الخطأ يحتمل ظهوره؟ وأي من نوعيْ الخطأ يعتبر أهم بالنسبة للمجتمع.

(٢٢) إن كانت قيمة $\bar{X} = 82$ و $\sigma = 16$ و $n = 100$ اختبر الفرض بأن $\mu = 86$ عند مستوى معنوية مناسب.

(٢٣) إن كانت قيمة $\bar{X} = 82$ و $\sigma = 26$ و $n = 25$ اختبر الفرض بأن $\mu = 86$ عند مستوى معنوية 0.05 .

(٢٤) خبرة سنوات عدة في امتحان اللغة الإنجليزية لدخول الجامعة أعطي متوسط النقاط المحققة مساويًا للقيمة 64 نقطة بانحراف معياري قدره 8 درجات. وقد حصل طلّاب مدينة ما وعددهم 55 على وسط نقاط قدره 68 نقطة فهل يمكن التحقق من أن هؤلاء الطلّاب أحسن مستوى في الإنجليزية من بقية الطلّاب؟ عند مستوى معنوية 0.05 .

(٢٥) يدّعي صانع سنّارات سمك أن اختباره للسنّارات ستعطي 8 رطل عند الاختبار فهل هو محق في دعواه؟ إن كانت عيّنة حجمها 50 تعطى $\bar{X} = 7$ أرطال و $S = 1.4$ رطل.

(٢٦) في عيّنة مكونة من 300 من الرجال و 500 من النساء الذين شاهدوا برنامجًا تلفزيونيًا معيّنًا وذكر 60 من الرجال و 200 من النساء أنهم يفضّلون هذا البرنامج.

أوجد: (i) 95 %(ii) 99 % حدود ثقة للفرق بين نسبة كل من الرجال ونسبة كل من النساء الذين شاهدوا هذا البرنامج ويفضّلونه.

(٢٧) يحتوي وعاء على عدد غير معروف من البِلْي الأحمر والأبيض. أخذت عيّنة عشوائية من 60 من البِلْي أختيرت مع الإرجاع من الوعاء أظهرت 70 % من البِلْي الأحمر.

أوجد (i) 90 % (ii) 99 % حدود ثقة للنسبة الفعلية للبِلْي الأحمر.

(٢٨) يدّعي منتج أن 95 % على الأقل من المعدّات التي يمد بها مصنعًا مطابقة للمواصفات. تم اختيار عيّنة من 200 وحدة من المعدّات ووجد أن بها 18 وحدة تالفة.

اختبر ادّعاء المنتج عند مستوى المعنوية (i) 0.01 (ii) 0.05 .

(٢٩) نسبة الطلاّب الذين حصلوا على تقدير A في مادة الفيزياء في إحدى الجامعات خلال فترة طويلة من الزمن كانت 10 % ، خلال فصل دراسي معيّن حصل 40 طالبًا على تقدير A من مجموعة من 300 طالب.

اختبر معنوية هذه النتيجة عند مستوى معنوية (i) 0.05 (ii) 0.01 .

(٣٠) عيّنة عشوائية من 300 مسمار من إنتاج A و 200 مسمار من إنتاج B وجد أن 28 مسمارًا من إنتاج A و 15 مسمارًا من إنتاج B تالفًا.

اختبر الفرض القائل أن:

i) هناك اختلاف في أداء الماكينتين.

ii) الماكينة B تعمل بصورة أفضل من الماكينة A .

استخدم مستوى للمعنوية (ا) 0.01 (ب) 0.05 .

الجـــداول

- جدول رقم (١) التوزيع الطبيعي $\Phi(z)$
- جدول رقم (٢) الأرقام العشوائية
- جدول رقم (٣) توزيع t (ت)

جدول رقم (١) التوزيع الطبيعي $\Phi(z)$

The cumulative standardized normal distribution function [Tabulated value = $Pr(Z < z_p) = P$]

z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
− .0	5000	4960	4920	4880	4840	4801	4761	4721	4681	4641
− .1	4602	4562	4522	4483	4443	4404	4364	4325	4286	4247
− .2	4207	4168	4129	4090	4052	4013	3974	3936	3897	3859
− .3	3821	3783	3745	3707	3669	3632	3594	3557	3520	3483
− .4	3446	3409	3372	3336	3300	3264	3228	3192	3156	3121
− .5	3085	3050	3015	2981	2946	2912	2877	2843	2810	2776
− .6	2743	2709	2676	2643	2611	2578	2546	2514	2483	2451
− .7	2420	2389	2358	2327	2297	2266	2236	2206	2177	2148
− .8	2119	2090	2061	2063	2005	1977	1949	1922	1894	1867
− .9	1841	1814	1788	1762	1736	1711	1685	1660	1635	1611
−1.0	1587	1562	1539	1515	1492	1469	1446	1423	1401	1379
−1.1	1357	1335	1314	1292	1271	1251	1230	1210	1190	1170
−1.2	1151	1131	1112	1093	1075	1056	1038	1020	1003	09853
−1.3	09680	09510	09342	09176	09012	08851	08691	08534	08379	08226
−1.4	08076	07927	07780	07636	07493	07353	07215	07078	06944	06811
−1.5	06681	06552	06426	06301	06178	06057	05938	05821	05705	05592
−1.6	05480	05370	05262	05155	05050	04947	04846	04746	04648	04551
−1.7	04457	04363	04272	04182	04093	04006	03920	03836	03754	03673
−1.8	03593	03515	03438	03362	03288	03216	03144	03074	03005	02938
−1.9	02872	02807	02743	02680	02619	02559	02500	02442	02385	02330
−2.0	02275	02222	02169	02118	02068	02018	01970	01923	01876	01831
−2.1	01786	01743	01700	01659	01616	01578	01539	01500	01463	01426
−2.2	01390	01355	01321	01287	01255	01222	01191	01160	01130	01101
−2.3	01072	01044	01017	$0^{2}9903$	$0^{2}9642$	$0^{2}9387$	$0^{2}9137$	$0^{2}8894$	$0^{2}8656$	$0^{2}8424$
−2.4	$0^{2}8198$	$0^{2}7976$	$0^{2}7760$	$0^{2}7549$	$0^{2}7344$	$0^{2}7143$	$0^{2}6947$	$0^{2}6756$	$0^{2}6569$	$0^{2}6387$
−2.5	$0^{2}6210$	$0^{2}6037$	$0^{2}5868$	$0^{2}5703$	$0^{2}5543$	$0^{2}5386$	$0^{2}5234$	$0^{2}5085$	$0^{2}4940$	$0^{2}4799$
−2.6	$0^{2}4661$	$0^{2}4527$	$0^{2}4396$	$0^{2}4269$	$0^{2}4145$	$0^{2}4025$	$0^{2}3907$	$0^{2}3793$	$0^{2}3681$	$0^{2}3573$
−2.7	$0^{2}3467$	$0^{2}3364$	$0^{2}3264$	$0^{2}3167$	$0^{2}3072$	$0^{2}2980$	$0^{2}2890$	$0^{2}2803$	$0^{2}2718$	$0^{2}2635$
−2.8	$0^{2}2555$	$0^{2}2477$	$0^{2}2401$	$0^{2}2327$	$0^{2}2256$	$0^{2}2186$	$0^{2}2118$	$0^{2}2052$	$0^{2}1988$	$0^{2}1926$
−2.9	$0^{2}1866$	$0^{2}1807$	$0^{2}1750$	$0^{2}1695$	$0^{2}1641$	$0^{2}1589$	$0^{2}1538$	$0^{2}1489$	$0^{2}1441$	$0^{2}1395$
−3.0	$0^{2}1350$	$0^{2}1306$	$0^{2}1264$	$0^{2}1223$	$0^{2}1183$	$0^{2}1114$	$0^{2}1107$	$0^{2}1070$	$0^{2}1035$	$0^{2}1001$
.0	5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
.1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
.2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
.3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
.4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
.5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
.6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
.7	7580	7611	7642	7673	7703	7734	7764	7794	7823	7852
.8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
.9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1.0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8661
1.1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1.2	8849	8869	8888	8907	8925	8944	8962	8980	8997	90147
1.3	90320	90490	90658	90824	90988	91149	91309	91466	91621	91774
1.4	91924	92073	92220	92364	92507	92647	92785	92922	93056	93189
1.5	93319	93448	93574	93669	93822	93943	94062	94179	94295	94408
1.6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1.7	95543	95637	95728	95818	95907	95994	96080	96164	96246	96327
1.8	96407	96485	96562	96638	96712	96784	96856	96926	96995	97062
1.9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97670
2.0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
2.1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2.2	98610	98645	98679	98713	98745	98778	98809	98840	98870	98899
2.3	98928	98956	98983	$9^{2}0097$	$9^{2}0358$	$9^{2}0613$	$9^{2}0863$	$9^{2}1106$	$9^{2}1344$	$9^{2}1576$
2.4	$9^{2}1802$	$9^{2}2024$	$9^{2}2240$	$9^{2}2451$	$9^{2}2656$	$9^{2}2857$	$9^{2}3053$	$9^{2}3244$	$9^{2}3431$	$9^{2}3613$
2.5	$9^{2}3790$	$9^{2}3963$	$9^{2}4132$	$9^{2}4297$	$9^{2}4457$	$9^{2}4614$	$9^{2}4766$	$9^{2}4915$	$9^{2}5060$	$9^{2}5201$
2.6	$9^{2}5339$	$9^{2}5473$	$9^{2}5604$	$9^{2}5731$	$9^{2}5855$	$9^{2}5975$	$9^{2}6093$	$9^{2}6207$	$9^{2}6319$	$9^{2}6427$
2.7	$9^{2}6533$	$9^{2}6636$	$9^{2}6736$	$9^{2}6833$	$9^{2}6928$	$9^{2}7020$	$9^{2}7110$	$9^{2}7197$	$9^{2}7282$	$9^{2}7365$
2.8	$9^{2}7445$	$9^{2}7523$	$9^{2}7599$	$9^{2}7673$	$9^{2}7744$	$9^{2}7814$	$9^{2}7882$	$9^{2}7948$	$9^{2}8012$	$9^{2}8074$
2.9	$9^{2}8134$	$9^{2}8193$	$9^{2}8250$	$9^{2}8305$	$9^{2}8359$	$9^{2}8411$	$9^{2}8462$	$9^{2}8511$	$9^{2}8559$	$9^{2}8605$
3.0	$9^{2}8650$	$9^{2}8694$	$9^{2}8736$	$9^{2}8777$	$9^{2}8817$	$9^{2}8856$	$9^{2}8893$	$9^{2}8930$	$9^{2}8965$	$9^{2}8999$

* Reprinted with permission from A. Hald, "*Statistical Tables and Formulas*," John Wiley & Sons, Inc., New York, 1952.

Note: $.0^{2}1350 = .001350$ $.9^{2}8650 = .998650$

جدول رقم (٢) الأرقام العشوائية

48611	62866	33963	14045	79451	04934	45576
78812	03509	78673	73181	29973	18664	04555
19472	63971	37271	31445	49019	49405	46925
51266	11569	08697	91120	64156	40365	74297
55806	96275	26130	47949	14877	69594	83041
77527	81360	18180	97421	55541	90275	18213
77680	58788	33016	61173	93049	04694	43534
15404	96554	88265	34537	38526	67924	40474
14045	22917	60718	66487	46346	30949	03173
68376	43918	77653	04127	69930	43283	35766
93385	13421	67957	20384	58731	53396	59723
09858	52104	32014	53115	03727	98624	84616
93307	34116	49516	42148	57740	31198	70336
04794	01534	92058	03157	91758	80611	45357
86265	49096	97021	92582	61422	75890	86442
65943	79232	45702	67055	39024	57383	44424
90038	94209	04055	27393	61517	23002	96560
97283	95943	78363	36498	40662	94188	18202
21913	72958	75637	99936	58715	07943	23748
41161	37341	81838	19389	80336	46346	91895
23777	98392	31417	98547	92058	02277	50315
59973	08144	61070	73094	27059	69181	55623
82690	74099	77885	23813	10054	11900	44653
83854	24715	48866	65745	31131	47636	45137
61980	34997	41825	11623	07320	15003	56774
99915	45821	97702	87125	44488	77613	56823
48293	86847	43186	42951	37804	85129	28993
33225	31280	41232	34750	91097	60752	69783
06846	32828	24425	30249	78801	26977	92074
32671	45587	79620	84831	38156	74211	82752
82096	21913	75544	55228	89796	05694	91552
51666	10433	10945	55306	78562	89630	41230
54044	67942	24145	42294	27427	84875	37022
66738	60184	75679	38120	17640	36242	99357
55064	17427	89180	74018	44865	53197	74810
69599	60264	84549	78007	88450	06488	72274
64756	87759	92354	78694	63638	80939	98644
80817	74533	68407	55862	32476	19326	95558
39847	96884	84657	33697	39578	90197	80532
90401	41700	95510	61166	33757	23279	85523
78227	90110	81378	96659	37008	04050	04228
87240	52716	87697	79433	16336	52862	69149
08486	10951	26832	39763	02485	71688	90936
39338	32169	03713	93510	61244	73774	01245
21188	01850	69689	49426	49128	14660	14143
13287	82531	04388	64693	11934	35051	68576
53609	04001	19648	14053	49623	10840	31915
87900	36194	31567	53506	34304	39910	79630
81641	00496	36058	75899	46620	70024	88753
19512	50277	71508	20116	79520	06269	74173

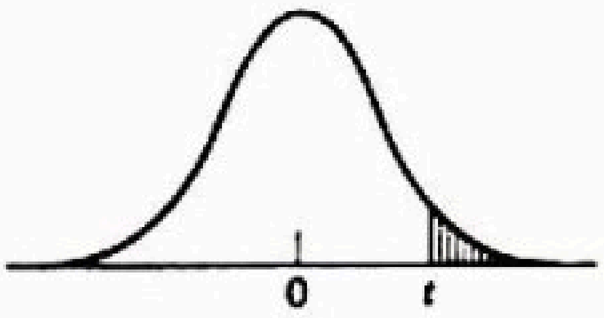
تابع جدول رقم (٢) الأرقام العشوائية

04433	80674	24520	18222	10610	05794	37515
60298	47829	72648	37414	75755	04717	29899
67884	59651	67533	68123	17730	95862	08034
89512	32155	51906	61662	64130	16688	37275
32653	01895	12506	88535	36553	23757	34209
95913	15405	13772	76638	48423	25018	99041
55864	21694	13122	44115	01601	50541	00147
35334	49810	91601	40617	72876	33967	73830
57729	32196	76487	11622	96297	24160	09903
86648	13697	63677	70119	94739	25875	38829
30574	47609	07967	32422	76791	39725	53711
81307	43694	83580	79974	45929	85113	72268
02410	54905	79007	54939	21410	86980	91772
18969	75274	52233	62319	08598	09066	95288
87863	82384	66860	62297	80198	19347	73234
68397	71708	15438	62311	72844	60203	46412
28529	54447	58729	10854	99058	18260	38765
44285	06372	15867	70418	57012	72122	36634
86299	83430	33571	23309	57040	29285	67870
84842	68668	90894	61658	15001	94055	36308
56970	83609	52098	04184	54967	72938	56834
83125	71257	60490	44369	66130	72936	69848
55503	52423	02464	26141	68779	66388	75242
47019	76273	33203	29608	54553	25971	69573
84828	32592	79526	29554	84580	37859	28504
68921	08141	79227	05748	51276	57143	31926
36458	96045	30424	98420	72925	40729	22337
95752	59445	36847	87729	81679	59126	59437
26768	47323	58454	56958	20575	76746	49878
42613	37056	43636	58085	06766	60227	96414
95457	30566	65482	25596	02678	54592	63607
95276	17894	63564	95958	39750	64379	46059
66954	52324	64776	92345	95110	59448	77249
17457	18481	14113	62462	02798	54977	48349
03704	36872	83214	59337	01695	60666	97410
21538	86497	33210	60337	27976	70661	08250
57178	67619	98310	70348	11317	71623	55510
31048	97558	94953	55866	96283	46620	52087
69799	55380	16498	80733	96422	58078	99643
90595	61867	59231	17772	67831	33317	00520
33570	04981	98939	78784	09977	29398	93896
15340	93460	57477	13898	48431	72936	78160
64079	42483	36512	56186	99098	48850	72527
63491	05546	67118	62063	74958	20946	28147
92003	63868	41034	28260	79708	00770	88643
52360	46658	66511	04172	73085	11795	52594
74622	12142	68355	65635	21828	39539	18988
04157	50079	61343	64315	70836	82857	35335
86003	60070	66241	32836	27573	11479	94114
41268	80187	20351	09636	84668	42486	71303

* Based on parts of *Tables of 105,000 Random Decimal Digits*. Interstate Commerce Commission, Bureau of Transport Economics and Statistics, Washington D.C.

جدول رقم (٣) توزيع أستودينت t (ت)

العمود الأول خاص بدرجات الحرية (ν) . عناوين بقية الأعمدة تعطي الاحتمالات (α) لقيمة t التي تزيد عن قيمة المدخل. استخدام التماثل لقيمة t السالبة.



ν \ α	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

المصدر: عن (بول. ج هويل، ١٩٨٤م)

المراجــع

- المراجع العربية
- المراجع الأجنبية

المراجع العربية

أبوصالح، محمد صبحي و عوض، عدنان محمد (١٩٨٣م). *مقدمة في الاحصاء*. نيويورك: دار جون وايلي وأبنائه.

الصياد، جلال و سمرة، عادل (١٩٧٦م). *مبادىء الاحصاء لطلاب الدراسات الأدبية*، الطبعة الأولى. المملكة العربية السعودية، جدة: جامعة الملك عبدالعزيز.

زايد، مصطفى (١٩٨٤م). *الاحصاء ووصف البيانات*. المملكة العربية السعودية، الرياض: دار العلوم للطباعة والنشر.

سرحان، أحمد عبادة (١٩٦٥م). *طرق التحليل الاحصائي*. القاهرة: دار المعارف.

عاشـور، سمير كامل (١٩٧٧م). *مبادىء الاحصاء الوصفي والتحليلي*. القاهرة: معهد الإحصاء.

ليبشتــز، سيمــور (١٩٧٤م). *سلسلة ملخصــات شوم في الاحتــالات*. لنــدن: ماكجروهيل للنشر؛ ترجمة: **سامح داود** ومراجعة: **عبدالعظيم أنيس**. المملكة العربية السعودية، الرياض: دار المريخ.

ليبشتــز، سيمــور (١٩٧٤م). *سلسلة ملخصــات شوم في الاحصــاء*. لنــدن: ماكجـروهيل للنشر؛ ترجـمـة: **شعبـان عبـدالحميـد شعبـان** ومـراجعة: **أحمد الموازيني**. المملكة العربية السعودية، الرياض: دار المريخ.

مصطفى، مدني دسوقي (١٩٧٥م). *مبادىء في علم الاحصاء*. القاهرة: دار النهضة العربية.

مصطفى، مدني دسوقي (١٩٧٩م). *مبادىء في نظرية الاحتمالات والاحصاء الرياضي*. القاهرة: دار النهضة العربية.

هويل، كبول ج (١٩٨٤م). *المبادىء الأولية في الاحصاء*، الطبعة الثانية؛ ترجمة: بدرية شوقي عبدالوهاب ومحمد كامل الشربيني. نيويورك: دار جون وايلي وأبنائه.

المراجع الأجنبية

Bowen, Earl K. and Starr, Martin K. (1980). *Basic Statistics for Business and Economics*. London: McGrow-Hill International Company.

Ellfson, Kirk W.; Runyon, Richard P. and Haber, Audrey (1980). *Fundamentals of Social Statistics*. London: Addison-Wesley Publishing Company.

Ferguson, George A. (1981). *Statistical Analysis in Psychology and Education*. London: McGrow-Hill International Book Company.

Hassett, Weiss (1982). *Introductory Statistics*. London: Addison-Wesley Publishing Company.

John, Freund E. (1979). *Modern Elementery Statistics, 5th* Ed. Englewood Cliffs, New Jersey: Prentice Hall International Editions.

Lapin, Lawrence (1980). *Statistics Meaning and Methods*. New York: Harcourt Brace Jovanovich, Inc.

Mendenhall, William (1980). *Introduction to Probability and Statistics*. North Scituate: Duxbury Press.

ثبت المصطلحات

- عربي - إنجليزي
- إنجليزي - عربي

عربي - إنجليزي

ا

اتحاد	Union (U)
احتمال حدوث الحادث A	Probability of A ($P_{(A)}$)
احتمال حدوث الحادثة A بشرط حدوث الحادثة B	Conditional probability ($P_{(A/B)}$)
الانحراف المتوسط	Mean deviation (M.D)
الانحراف المعياري للعيّنة	Sample standard deviation (s)
الانحراف المعياري للمجتمع	Population standard deviation (σ)

ت

التباديل	Permutation (^nP_r)
تباين العيّنة	Sample variance (S^2)
تباين المتغيّر العشوائي X	Variance of X [V(X)]
تباين المتوسط	Variance of $\bar{X}$ ($\sigma^2_{\bar{X}}$)
تباين المجتمع	Population variance (σ^2)

التفلطح	Kurtosis (K)
التقاطع	Intersection ($\cap$)
التكرار	Frequency (f)
التوافيق	Combination [^nC_r أو $\binom{n}{r}$]
توزيع برنولي (محاولة برنولي)	Bernulli distribution
توزيع بواسون	Poisson distribution
توزيع ت (t)	T-distribution
توزيع ذي الحدين (الثنائي)	Binomial distribution
التوزيع الطبيعي	Normal distribution
التوزيع الطبيعي القياسي	Standard normal distribution
توزيع فوق الهندسي	Hypergeometric distribution
التوزيع المنتظم المتقطّع	Discrete uniform distribution
توقع المتغيّر العشوائي X	Expectation of X [E(X)]
توقع المتوسط	Expectation of $\bar{X}$ [$M_{\bar{X}}$]

ج

جبرسجما	σ-algebra (ξ)

خ

الخطأ المعياري	Standard error ($\sigma_{\bar{X}}$)

د

دالة التوزيع التراكمي للتوزيع الطبيعي القياسي . . .	Cumulative standard normal distribution function $\Phi(x)$
دالة التوزيع التراكمي للمتغيّر العشوائي X	Cumulative distribution function [F(.)]

Probability mass function [f(.)]	دالة الكتلة الاحتمالية للمتغيّر العشوائي المنفصل X
Probability density function [f(.)]	دالة الكثافة الاحتمالية للمتغيّر العشوائي المتّصل X
Density function of standard normal $\varnothing(.)$	دالة الكثافة للتوزيع الطبيعي القياسي

ر

First quartile (Q_1)	الربيع الأول (الأدنى)
Third quartile (Q_3)	الربيع الثالث (الأعلى)
Quartiles (Q_i)	الربيعيـات
Summation (Σ)	رمز التجميع

ط

Class length (L)	طول الفئة

ع

Sample points [n(S)]	عدد نقاط فراغ العيّنة
Deciles (D_i)	العشيرات

ف

Sample space (S)	فراغ العيّنة
Alternative hypothesis (H_1)	فرض البديل
Null hypothesis (H_0)	فرض العدم
Classes of sets (ξ)	فصول المجموعات الجزئية

ق

Maximum error of estimate (E)	القيمة العظمى للخطأ في التقدير
Co-domain [X(S)]	القيم الممكنة للمتغيّر العشوائي X (المجال المقابل)

م

المئينـات	Percentiles (P_i)
المتغيّر المعياري	Standardized variable (Z)
مجموعات حاصل الضرب	Cartesian product ($A \times B$)
مجموعة جزئية	Subset ($\subset$)
مجموعة خالية	Empty set (ϕ)
مجموعة القوى	Power set (P_A)
مجموعة ليست جزئية	Not subset ($\not\subset$)
مجموعة مكملة	Complement set (A^c) أو ($\bar{A}$)
المـــدى	Range (R)
مركز الفئة	Class mark (x)
مستوى المعنوية	Level of significance (α)
مضروب العدد ن (لـــن)	Factional n (n!)
معامل الاختلاف	Coefficient of variation (C.V)
معامل الارتباط الخطي	Coefficient of linear correlation (r)
معامل ارتباط الرتب	Coefficient of rank correlation (r_s)
معامل الاقتران والتوافق	Coefficient of contingency (c.c)
معامل الالتواء	Skewness coefficient (v)
معامل الانحدار	Coefficient regression (b)
المنـــوال	Mode (Mod)

ن

نصف المدى الربيعي	Semi-interquartile range (Q)

و

الوسط التوافقي	Harmonic mean (H)

Arithmetic mean ($\bar{x}$)	الوسط الحسابي
Weighted mean ($\bar{x}_w$)	الوسط المرجّح
Geometric mean (G.M)	الوسط الهندسي
Median (Med)	الوسيــط

إنجليزي - عربي

A

Alternative hypothesis $[H_1]$	فرض البديل
Arithmetic mean $[\bar{x}]$	الوسط الحسابي

C

Cartesian product of A and B	مجموعات حاصل الضرب $(A \times B)$
Classes of sets (ξ)	فصول المجموعات الجزئية
Class length (L)	طـول الفئـة
Class mark (x)	مركز الفئـة
Co-domain of a random variable X $[X(S)]$	المجال المقابل للمتغيّر العشوائي X
Coefficient of contingency (c.c)	معامل الاقتران والتوافق
Coefficient of linear correlation (r)	معامل الارتباط الخطي
Coefficient of rank correlation (r_s)	معامل ارتباط الرتب
Coefficient of regression (b)	معامل الانحدار
Coefficient of variation (C.V)	معامل الاختلاف
Combination of n taken r $[{}^nC_r$ or $\binom{n}{r}]$	التوافيق
Complement set of A $[A^c$ or $\bar{A}]$	مجموعة مكمّلة

Conditional probability [P(A/B)]	احتمال حدوث A بشرط حدوث B
Cumulative distribution function [F(.)]	دالة التوزيع التراكمي
Cumulative standard normal distribution [Φ(.)]	دالة التوزيع التراكمي للمتغيّر العشوائي الطبيعي القياسي

D

Deciles [D_i]	العشيرات
Density function of standard normal [ϕ(.)]	دالة الكثافة للتوزيع الطبيعي القياسي

E

Empty set [ϕ]	المجموعة الخالية
Expectation of X [E(X)]	توقع المتغيّر العشوائي X
Expectation of $\bar{X}$ [$\mu_{\bar{X}}$]	توقع المتوسط $\bar{X}$

F

Factorial n [n!]	مضروب ـ ن [لـ ن]
First quartile [Q_1]	الربيع الأول (الأدنى)
Frequency [f]	التكـــرار

G

Geometric mean [G.M]	الوسط الهندسي

H

Harmonic mean [H]	الوسط التوافقي

I

Intersection [$\cap$]	التقاطـع

K

Kurtosis [K]	التفلطـح

L

Level of significance [α]	مستوى المعنوية

M

Maximum error of estimate (E)	القيمة العظمى للخطأ في التقدير
Mean deviation [M.D]	الانحراف المتوسط
Median [Med]	الوسيـط
Mode [Mod]	المنـوال

N

Not subset [$\not\subset$]	مجموعة ليست جزئية
Null hypothesis [H_0]	فـرض العدم

P

Percentiles [P_i]	المئينـات
Permutation of n taken r [${}^{n}p_r$]	التباديل
Population standard deviation [σ]	الانحراف المعياري للمجتمع
Population variance [σ^2]	التباين للمجتمع
Power set of A [P_A]	مجموعة القـوى
Probability density function [$f_X(\cdot)$]	دالة الكثافة الاحتمالية للمتغيّر العشوائي X

Probability mass function $[f_X(\cdot)]$	دالة الكتلة الاحتمالية للمتغيّر العشوائي X
Probability of A $[P(A)_{(A)}]$	احتمال حدوث الحادثة A

Q

Quartiles $[Q_i]$	الربيعيات

R

Range [R]	المـــدى

S

Sample points [n(S)]	عدد نقاط فراغ العيّنة
Sample space [S]	فراغ العيّنة
Sample standard deviation [S]	الإنحراف المعياري للعيّنة
Sample variance	التباين للعيّنة
Semi-interquartile range [Q]	نصف المدى الربيعي
Skewness [v]	معامل الإلتواء
Standard error $[\sigma_{\overline{X}}]$	الخطأ المعياري
Standardized variable [Z]	المتغير العشوائي المعياري
Subset $[\subset]$	مجموعة جزئية
Summation $[\Sigma]$	رمـز التجميع

T

T-distribution	توزيع ت (t)
Third quartile $[Q_3]$	الربيع الثالث

U

Union [U]	الاتحـــاد

V

Variance of X [V(X)]	تباين X
Variance of $\bar{X}$ [$\sigma^2_{\bar{X}}$]	تباين $\bar{X}$

W

Weighted mean [$\bar{x}_w$]	الوسط المرجّح

كشاف الموضوعات

د

ر

ط

ع

ف

ق

م

ردمك: ٩٩٦٠-٠٥-٥٦٩-٨

ISBN: 9960-05-569-8